
СЕКЦИЯ 10 | **НЕЧЕТКИЕ МОДЕЛИ
И МЯГКИЕ ВЫЧИСЛЕНИЯ**

УДК 004.891.3

**АЛГОРИТМ ПОИСКА АНОМАЛИЙ В ПРОЦЕССАХ НА
ОСНОВЕ НЕЧЕТКИХ ТЕНДЕНЦИЙ ВРЕМЕННЫХ РЯДОВ**

Т.В. Афанасьева (*tv.afanasjeva@gmail.com*)

Д.В. Заварзин (*dzavarzin91@gmail.com*)

Ульяновский государственный технический университет, Ульяновск

В работе предлагается новая классификация возможных аномалий в процессах, параметры которых представимы временными рядами. Особенностью классификации аномалий является расширение традиционной классификации за счет нечетких локальных тенденций. Формализованы условия обнаружения введенных классов аномалий на временных рядах и предложена схема программно-реализованного алгоритма.

Введение

Задача поиска и обнаружения аномалий в процессах является актуальной в широком диапазоне сфер применения: экология, электронная коммерция, анализ технологических процессов, надежность технических и информационных систем, веб-аналитика, выявление знаний и интеллектуальный анализ данных [Chandola, 2009], [Cheboli, 2010].

При решении данной задачи с целью диагностики процессов, интерпретированных временными рядами (ВР), целесообразно применять методы, основанные на поиске аномалий в поведении ВР. Так как ВР обычно отображаются графиками, то процесс обнаружения аномалий обычно заключается в наблюдательной дея-

тельности эксперта по обнаружению типичных или нетипичных значений или тенденций.

Обычно задача обнаружения аномалий на временных рядах (ВР) возникает в связи с пропуском значений или появлением значений, выходящих за пределы диапазона допустимого интервала. Однако эти подходы не учитывают аномалии, связанные с нарушением поведения ВР, выраженные в изменении локальных тенденций.

В данной работе предлагается алгоритм поиска и обнаружения аномалий в процессах, протекающих в условиях неопределенности, на основе анализа нечетких локальных тенденций. Результатом работы являются новые знания о возникающих аномальных тенденциях, которые невозможно определить другими методами.

Структура статьи следующая. В первом разделе рассматриваются задача поиска и обнаружения аномалий на ВР, а также используемые при этом алгоритмы. Во втором разделе мы рассмотрим классификацию аномалий на ВР. В третьем разделе описаны алгоритм поиска и обнаружения аномалий на ВР с учетом аномалий в нечетких локальных тенденциях и его программная реализация.

1. Задача поиска и обнаружения аномалий на временных рядах

Согласно обзору [Cheboli, 2010], поиск и обнаружение аномалий на ВР рассматриваются как совокупность задач обнаружения аномалий на всем ВР, на участке ВР или в отдельной точке ВР.

Например, в работе [Salvador, 2004] предложен алгоритм поиска и обнаружения аномалий на ВР, который успешно внедрен и применяется в процессе экспертного анализа параметров космических челноков NASA. В этой работе авторы решают задачу обнаружения аномалий ВР с помощью применения алгоритма кластеризации Gеско для выявления аномальных состояний, описанных логическими правилами (алгоритм классификации RIPPER). Кроме этого, подсчитывается количество аномальных состояний, которое не должно превышать заданный исследователем порог. Исследования показали, что метод сравним по эффективности работы с анализом данных человеком-экспертом.

В работе [Wei, 2005] все множество значений ВР разбивается на несколько уровней, каждому из которых сопоставляется символ заданного конечного алфавита. В этом случае для поиска и обнару-

ружения аномалий используется скользящее окно и сравнение с ВР, характеризующим нормальное поведение процесса. Данный алгоритм использовался для анализа ВР данных ЭКГ, мониторинга телеметрии космического челнока, видеопотоков и т.д.

В работах [Cheboli, 2010], [Ковалев, 2013] приведены модели и алгоритмы поиска и обнаружения аномалий на ВР, такие как:

1. Статистические модели (AR, ARIMA, фильтры Калмана).
2. Нечеткие модели (алгоритмы FRCM, SSOD, FUCOT).
3. Скрытые Марковские процессы.
4. Кластеризация и классификация (алгоритм кластеризации с-средних, k-средних).
5. Метод скользящих окон.

2. Классификация аномалий на временных рядах

Временной ряд – это последовательность дискретных упорядоченных в неслучайные равноотстоящие моменты времени измерений (показателей наблюдений) $X = \{x_t\}$, где $t = \overline{1, n}$, характеризующих уровни состояния изучаемого процесса. Практически любые наблюдения, полученные в результате мониторинга процессов в различных сферах человеческой деятельности (безопасность, медицина, экология и т.д.), могут быть представлены в виде некоторой последовательности, зависимой от времени, то есть в форме ВР.

Рассмотрим различные уровни представления ВР в рамках теории нечетких систем. Нечеткий временной ряд (НВР) [Ярушкина, 2007] – это последовательность нечетких меток, зафиксированных в равноотстоящие дискретные временные отсчеты $\tilde{X} = \{\tilde{x}_t\}$, где $t = \overline{1, n}$. Для НВР было формализовано понятие нечеткой тенденции, выражающей качественное изменение на временном интервале НВР [Афанасьева, 2009].

Нечеткая тенденция (НТ) [Афанасьева, 2011] ВР $\tau \in \mathfrak{Z}$ НВР $\tilde{X} = \{\tilde{x}_t\}$, $t = \overline{1, n}$ может быть представлена структурно-лингвистической моделью в виде отношения, построенного на декартовом произведении свойств нечеткой тенденции $\tilde{V} \times \tilde{A} \times \Delta \tilde{T} \rightarrow \mathfrak{Z}$:

$$\tau = \langle \tilde{v}, \tilde{a}, \Delta \tilde{T}, \mu \rangle, \quad (2.1)$$

где \tilde{v} – функция принадлежности нечеткого множества типа НТ $\tilde{v} \in \tilde{V}$, выражающая направление изменения (например, $\tilde{v}_1 =$ «рост», $\tilde{v}_2 =$ «стабильность», $\tilde{v}_3 =$ «падение» и т.д.); \tilde{a} – функция

принадлежности нечеткого множества интенсивности НТ, $\tilde{a} \in \tilde{A}$, выраженное лингвистическим термом (например, $\tilde{a}_1 =$ «большое», $\tilde{a}_2 =$ «среднее», $\tilde{a}_3 =$ «малое» и т.д.); $\Delta\tilde{t}$ – продолжительность НТ, где $\Delta\tilde{t} \in \Delta\tilde{T}$; μ – функция принадлежности участка НВР, ограниченного интервалом $\Delta\tilde{t}$ НТ.

Введем классификацию аномалий по компонентам ВР $X = (x_t, \tilde{x}_t, \tilde{\tau}_t)$, где x_t – элементарное четкое наблюдение на ВР; \tilde{x}_t – элементарное нечеткое наблюдение на ВР; $\tilde{\tau}_t$ – локальная нечеткая тенденция, $\tau_t = \langle \tilde{v}_t, \tilde{a}_t, \Delta\tilde{t}_t, \mu \rangle$.

Определение 1.

Пусть $X = \{x_t\}$, $t = \overline{1, n}$ – дискретный ВР, наблюдения в котором выражены числовыми значениями и известно множество аномальных значений $X' = \{x'_1, x'_2, \dots, x'_n\}$. Тогда условие обнаружения аномального значения ВР $X = \{x_t\}$ I класса может быть выражено:

$$x_t = \begin{cases} x_t^{valid}, & x_t \notin X' \\ x_t^{invalid}, & x_t \in X' \end{cases} \quad (2.2)$$

где x_t^{valid} – допустимое значение величины x_t , $x_t^{invalid}$ – недопустимое значение величины x_t , $X' = \{x'_1, x'_2, \dots, x'_n\}$ – множество аномальных значений величины x_t на ВР X .

Определение 2.

Пусть $\tilde{X} = \{\tilde{x}_t\}$, $t = \overline{1, n}$ – дискретный нечеткий ВР, наблюдения в котором выражены нечеткими значениями.

Тогда условие обнаружения аномального нечеткого значения нечеткого ВР $\tilde{X} = \{\tilde{x}_t\}$ II класса может быть выражено:

$$\tilde{x}_t = \begin{cases} \tilde{x}_t^{valid}, & \tilde{x}_t \notin X \\ \tilde{x}_t^{invalid}, & \tilde{x}_t \in X \end{cases} \quad (2.3)$$

где \tilde{x}_t^{valid} – допустимое нечеткое значение величины \tilde{x}_t , $\tilde{x}_t^{invalid}$ – недопустимое нечеткое значение величины \tilde{x}_t , $\tilde{X}' = \{\tilde{x}'_1, \tilde{x}'_2, \dots, \tilde{x}'_n\}$ – множество аномальных нечетких значений величины \tilde{x}_t на ВР X .

Определение 3.

Пусть $\mathfrak{S} = \{\tilde{\tau}_t\}$, $t = \overline{1, m}$ – дискретный ВР, состоящий из m -локальных нечетких тенденций.

Тогда, согласно (2.1), условия обнаружения аномального поведения нечеткого ВР $\tilde{X} = \{\tilde{x}_t\}$ III класса по компонентам нечетких локальных тенденций могут быть выражены:

$$\tilde{v}_t = \begin{cases} \tilde{v}_t^{valid}, \tilde{v}_t \notin \tilde{V}' \\ \tilde{v}_t^{invalid}, \tilde{v}_t \in \tilde{V}' \end{cases} \quad (2.4)$$

$$\tilde{a}_t = \begin{cases} \tilde{a}_t^{valid}, \tilde{a}_t \notin \tilde{A}' \\ \tilde{a}_t^{invalid}, \tilde{a}_t \in \tilde{A}' \end{cases} \quad (2.5)$$

$$\Delta \tilde{t}_t = \begin{cases} \Delta \tilde{t}_t^{valid}, \Delta \tilde{t}_t \notin \Delta \tilde{T}' \\ \Delta \tilde{t}_t^{invalid}, \Delta \tilde{t}_t \in \Delta \tilde{T}' \end{cases} \quad (2.6)$$

где \tilde{v}_t^{valid} – допустимое значение типа НТ \tilde{v}_t , $\tilde{v}_t^{invalid}$ – недопустимое значение типа НТ \tilde{v}_t , $\tilde{V}' = \{\tilde{v}'_1, \tilde{v}'_2, \dots, \tilde{v}'_n\}$ – множество аномальных типов НТ на ВР X ;

\tilde{a}_t^{valid} – допустимое значение интенсивности НТ \tilde{a}_t , $\tilde{a}_t^{invalid}$ – недопустимое значение интенсивности НТ \tilde{a}_t , $\tilde{A}' = \{\tilde{a}'_1, \tilde{a}'_2, \dots, \tilde{a}'_n\}$ – множество аномальных интенсивностей НТ на ВР X ;

$\Delta \tilde{t}_t^{valid}$ – допустимое значение продолжительности НТ $\Delta \tilde{t}_t$, $\Delta \tilde{t}_t^{invalid}$ – недопустимое значение продолжительности НТ $\Delta \tilde{t}_t$, $\Delta \tilde{T}' = \{\Delta \tilde{t}'_1, \Delta \tilde{t}'_2, \dots, \Delta \tilde{t}'_n\}$ – множество аномальных значений продолжительности НТ на ВР X .

Однако в выражениях (2.2)–(2.6) рассматриваются аномалии для конкретных независимых случаев, когда аномальность определяется числовым значением, нечетким значением или компонентами нечеткой тенденции на ВР. Введем модель, в которой сочетаются эти три ранее рассмотренных случая, для того, чтобы определить новый класс аномальности. При этом для простоты в качестве параметра нечеткой тенденции будем руководствоваться только типом ее изменения – \tilde{v}_t .

Определение 4.

Пусть \tilde{v}_t^{pred} – характеризует тип предыдущей нечеткой тенденции ВР; \tilde{x}_t – нечеткая оценка текущего состояния нечеткого ВР; \tilde{v}_t^{post} – характеризует тип будущей (возможно, предсказанное) нечеткой тенденции нечеткого ВР. Тогда для обнаружения аномального поведения ВР, выраженного в аномальном совместном действии нечетких объектов нечеткого ВР, рассмотрим тройку нечетких объектов:

$$\tilde{G} = \{\tilde{v}_t^{post}, \tilde{x}_t, \tilde{v}_t^{pred}\}, \quad (2.7)$$

где $\tilde{G} = \{\tilde{g}_t\}$ – нечеткий временной ряд, образованный такими тройками. Пусть известно множество аномальных троек $\tilde{G}' = \{\tilde{g}'_1, \tilde{g}'_2, \dots, \tilde{g}'_n\}$.

Тогда условие обнаружения аномального поведения нечеткого ВР $\tilde{X} = \{\tilde{x}_t\}$ IV класса, выраженного в терминах троек (2.7), интегрирующего знания об отношении нечетких локальных тенденций и нечеткого значения ВР, может быть выражено в виде:

$$\tilde{g}_t = \begin{cases} \tilde{g}_t^{valid}, \tilde{g}_t \notin \tilde{G}' \\ \tilde{g}_t^{invalid}, \tilde{g}_t \in \tilde{G}' \end{cases} \quad (2.8)$$

где \tilde{g}_t^{valid} – допустимое значение тройки \tilde{g}_t , $\tilde{g}_t^{invalid}$ – недопустимое значение тройки \tilde{g}_t , $\tilde{G}' = \{\tilde{g}'_1, \tilde{g}'_2, \dots, \tilde{g}'_n\}$ – множество аномальных троек на ВР X .

Отношение, задаваемое тройками (2.7) между нечеткими локальными тенденциями, и условие (2.8) формализуем с помощью модели нечеткого логического вывода в виде множества нечетких правил «Если-То»:

P_1 : ЕСЛИ \tilde{v}_{i1}^{pred} есть A_{i1} ИЛИ \tilde{x}_{i1} есть B_{i1} ТО \tilde{g}_{i1} есть G_{i1}

P_2 : ЕСЛИ (\tilde{v}_{i2}^{pred} есть A_{i2} ИЛИ \tilde{x}_{i1} есть B_{i2}) И \tilde{v}_{i1}^{post} есть C_{i1} ТО \tilde{g}_{i2} есть G_{i2}

...

P_i : ЕСЛИ \tilde{v}_{ii}^{pred} есть A_{ij} И (\tilde{x}_{ii} есть B_{ij} ИЛИ \tilde{v}_{ii}^{post} есть C_{ij}) ТО \tilde{g}_{ii} есть G_{ij}

3. Описание алгоритма обнаружения аномалий

Разработан алгоритм, учитывающий введенную классификацию аномалий на ВР. Особенностью разработанного алгоритма является тот факт, что он позволяет обнаруживать аномалии не только в числовых значениях, но и нечетких локальных тенденциях ВР.

Схема работы алгоритма обнаружения аномальных значений на ВР представлена на рис. 1. Алгоритм состоит из 9 функциональных блоков:

1. Блок «INPUT» предназначен для ввода исходных данных (исходного ВР, параметров аномалий).

2. Блок «FUZZY» предназначен для формирования нечетких лингвистических термов и описывающих их функций принадлежности.

3. Блок «А-I» определяет аномалии I класса (согласно формуле (2.2)).

4. Блок «А-II» определяет аномалии II класса (согласно формуле (2.3)).

5. Блок «A-III» определяет аномалии III класса (согласно формуле (2.4)).

6. Блок «A-IV» определяет аномалии IV класса (согласно формуле (2.8)).

7. Блок «VIZUALIZATOR» строит графики временных рядов, функций принадлежности лингвистических термов и отображает их пользователю.

8. Блок «ТЕХТ» представляет результаты вычислений в строковом формате.

9. Блок «OUTPUT» выводит результаты графической визуализации и строкового описания пользователю.

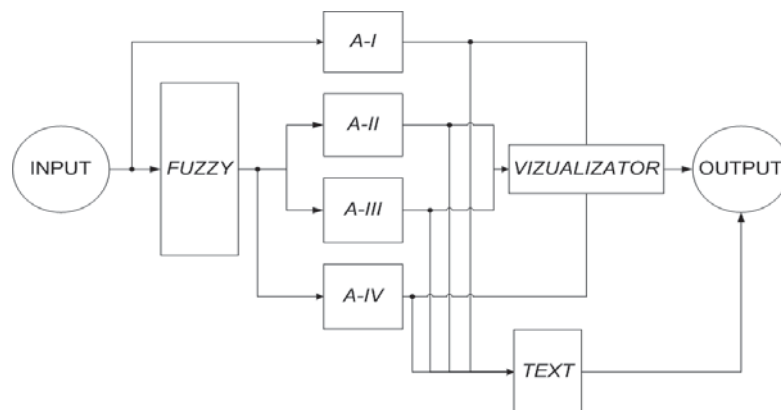


Рис. 1. Схема работы алгоритма поиска и обнаружения аномалий на ВР

Предложенный алгоритм показал свою эффективность в задаче обнаружения первых 3-х классов аномалий. В связи с ограниченностью статьи результаты его работы будут представлены в докладе.

Заключение

В настоящей работе был предложен алгоритм поиска и обнаружения аномалий в процессах, протекающих в условиях неопределенности, параметры которых представимы дискретными ВР. Алгоритм использует новую классификацию аномалий на ВР, особенностью которой является рассмотрение аномальных нечетких значений и локальных тенденций:

1. Аномалии числовых значений ВР.

2. Аномалии нечетких оценок числовых значений ВР.
3. Аномалии в нечетких тенденциях ВР.
4. Аномалии в отношениях соседних нечетких тенденций.

Направление будущей работы связано с проведением экспериментов и определением критериев качества обнаружения аномальных нечетких локальных тенденций.

СПИСОК ЛИТЕРАТУРЫ

[Афанасьева, 2009] Афанасьева Т.В., Ярушкина Н.Г. Нечеткое моделирование временных рядов и анализ нечетких тенденций. – Ульяновск: УлГТУ, 2009. – 299 с.

[Афанасьева, 2011] Афанасьева Т.В., Ярушкина Н. Г. Нечеткий динамический процесс с нечеткими тенденциями в анализе временных рядов // Вестник Ростовского государственного университета путей сообщения. – 2011. – №3. – С. 7–16.

[Ярушкина, 2007] Ярушкина Н.Г., Афанасьева Т.В. Нечеткие временные ряды как инструмент для оценки и измерения динамики процессов. // Журнал: Датчики и системы. 2007. № 12. – С. 46–50.

[Ярушкина, 2013] Ковалев С.М. Методы многошагового предсказания аномалий в темпоральных данных // Известия ЮФУ. Технические науки. Тематический выпуск ИНТЕЛЛЕКТУАЛЬНЫЕ САПР. №7, 2013. – С. 185–181.

[Chandola, 2009] Chandola V. Anomaly Detection : A Survey. – The University Of Minnesota. – 2009. – 72 с. – Режим доступа: <http://cucis.ece.northwestern.edu/projects/DMS/publications/AnomalyDetection.pdf> (дата обращения: 19.04.2014).

[Cheboli, 2010] Deepthi Cheboli. Anomaly Detection of Time Series. Faculty Of The Graduate School Of The University Of Minnesota. – 2010. – 75 с. – [Электронный ресурс] – Режим доступа: http://conservancy.umn.edu/bitstream/11299/92985/1/Cheboli_Deepthi_May2010.pdf (дата обращения: 20.04.2014).

[Salvador, 2004] Salvador St. Learning States and Rules for Time Series Anomaly Detection. – Applied Intelligence, Volume 23, Issue 3, pp. 241–255. – URL: <http://www.aaai.org/Papers/FLAIRS/2004/Flairs04-055.pdf> (дата обращения: 18.04.2014).

[Wei, 2005] Wei L., Kumar N. Assumption-Free Anomaly Detection in Time Series. – SSDBM'2005. – Proceedings of the 17th international conference on Scientific and statistical database management, pp. 237–240. – URL: <http://alumni.cs.ucr.edu/~ratana/SSDBM05.pdf> (дата обращения 19.04.2014).

УДК 004.827

МОДЕЛЬ НЕЧЕТКОГО СЦЕНАРИЯ РАЗВИТИЯ ВРЕМЕННОГО РЯДА

М.М. Зернов (*zmmioml@yandex.ru*)

филиал ФГБОУ ВПО «НИУ МЭИ» в г. Смоленске, Смоленск

В работе предлагается модель нечеткого сценария развития временного ряда, учитывающая неопределенность реализации сценария, а также позволяющая предсказать несколько этапов динамики временного ряда в соответствии с идентифицированным сценарием. Разработаны способы идентификации нечетких сценариев относительно фрагмента временного ряда.

Введение

В настоящее время не разработано достаточно гибкого математического аппарата для описания типовых сценариев развития временного ряда (ВР), инвариантных по отношению к масштабу по амплитуде и времени.

Подобные типовые сценарии, сохраняя общую структуру, реализуются каждый раз в виде уникальных фрагментов ВР, что приводит к необходимости введения неопределенности в их описание.

Имеющиеся нечеткие модели временного ряда или не вводят нечеткости во временную компоненту, что снижает обобщающую способность такой модели [Ярушкина и др., 2010], или используют ее как компоненту продукционных правил [Ковалев, 2007]. Недостатком последней модели является обязательное разделение нечеткого сценария на предпосылку и заключение (описывающее лишь один этап системной динамики), тогда как было бы полезно иметь возможность уже на ранних стадиях спрогнозировать несколько этапов развития ряда. Отсутствие инвариантности к масштабу как по амплитуде сигнала, так и по времени также является существенным недостатком.

1. Модель нечеткого сценария

Ниже предлагается модель нечеткого сценария, отличающаяся рассмотрением относительного нечеткого приращения сигнала за

нечеткое время, позволяющая сопоставлять сходные по структуре, но различные по временным характеристикам и/или масштабу отрезки временных рядов. Данная модель обобщает предложенную в [Зернов, 2014] модель для идентификации реализаций паттерна Эллиотта.

Нечеткий сценарий P образует последовательность нечетких волн (монотонных отрезков – стадий развития ряда):

$$P = (W_1, W_2, \dots, W_{Nw}),$$

где $W_i, i \in [1, Nw]$ – отдельная нечеткая волна/стадия:

$$W_i = \langle k_i, A_i, T_i \rangle,$$

где $k_i \in [0, i-1]$ – номер волны, базовой для рассматриваемой, A_i – нечеткая относительная амплитуда волны, T_i – нечеткая продолжительность волны.

Значения A_i и T_i представляют собой нечеткие множества на действительной оси и положительной полуоси соответственно:

$$A_i \in \mathbf{F}(\mathfrak{R}), T_i \in \mathbf{F}(\mathfrak{R}^+).$$

Практически без потери общности можем считать A_i и T_i нечеткими числами, т.е. нечеткими множествами на непрерывном действительном базовом множестве, обладающими свойствами выпуклости, компактности носителя и полунепрерывной сверху функцией принадлежности.

Волна описывает монотонное изменение амплитуды сигнала за указанное нечеткое время. При этом наличие локальных экстремумов полностью не исключается, но их локальность/область действия не должна превышать некоторой чувствительности. Нечеткая амплитуда является псевдолинейным приращением, указанным относительно амплитуды некоторой базовой волны. Иначе говоря, чтобы рассчитать абсолютную амплитуду волны W_i – A_i' , необходимо абсолютную величину базовой волны W_{k_i} умножить на величину A_i (расширенные арифметические операции над нечеткими числами будем обозначать как обычные):

$$A_i' = A_{k_i}' \cdot A_i.$$

Естественно, в сценарии должны присутствовать волны W_i , для

которых базовой волны нет ($k_i = 0$) – будем называть их основными, тогда:

$$A'_i = \begin{cases} A'_{k_i} \cdot A_i, k_i > 0, \\ A_i, k_i = 0. \end{cases}$$

Аналогичные соотношения вводятся и для нечеткой продолжительности T'_i .

В некоторых случаях нечеткие амплитуда и продолжительность могут базироваться на разных волнах, а величина k_i представлять собой двухмерный вектор – $k_i = (k_i^1, k_i^2)$. Тогда выражения для амплитуды и продолжительности записываются в виде:

$$A'_i = \begin{cases} A'_{k_i} \tilde{*} A_i, k_i^1 > 0, \\ A_i, k_i^1 = 0, \end{cases} \quad T'_i = \begin{cases} T'_{k_i} \tilde{*} T_i, k_i^2 > 0, \\ T_i, k_i^2 = 0. \end{cases}$$

В дальнейшем, если не оговорено иное, будем считать $k_i^1 = k_i^2 = k_i$.

Основные волны задают фактический масштаб сценария. В общем случае их можно считать переменными величинами, получающими означивание в процессе применения сценария. Неозначенное (нормированное) значение для нечеткой компоненты волны будем обозначать символом e , означающим четкое множество вида: $e = \{1/1\}$. Тогда такая основная волна W_i имеет вид: $W_i = \langle 0, e, e \rangle$. При подстановке в основную волну W_i ненормированного значения абсолютной амплитуды A'_i и/или продолжительности T'_i из сценария P получаем сценарий $P' = P(W_i := (A'_i, T'_i))$. Если более одной из основных волн получает означивание, значения для подстановки перечисляются через запятую.

2. Способы идентификации нечеткого сценария

Непосредственное применение модели нечеткого сценария описывается способами идентификации по представленному для анализа фрагменту временного ряда. Способы идентификации могут различаться в зависимости от возможности провести предварительную обработку временного ряда с целью выделить четкие реализации волн нечеткого сценария или провести означивание нечеткого сценария.

2.1. Способ идентификации означенного нечеткого сценария относительно фрагмента четкого временного ряда

Рассмотрим вначале вариант способа идентификации, при котором ряд не подвергается предварительной обработке (представлен набором дискретных значений), но сценарий означен.

Вначале нужно сформировать нечеткие временные интервалы волн, которые будут характеризовать, с какой уверенностью тот или иной момент времени относится к той или иной волне.

На основе абсолютной продолжительности волн $T_i' (i \in [1, N_w])$ определим нечеткое время окончания для каждой волны:

$$OT_i' = \sum_{j=1}^i T_j', j \in [1, N_w],$$

где знак суммы означает нечеткую арифметическую операцию суммирования значений абсолютных продолжительностей волн.

Будем рассматривать нечеткое время окончания 1-й волны T_1' как пересечение нечетких интервалов моментов времени, относящихся к 1-й I_1' и 2-й I_2' волне сценария (см. рис. 1.1) – $OT_1' = I_1' \cap I_2'$.

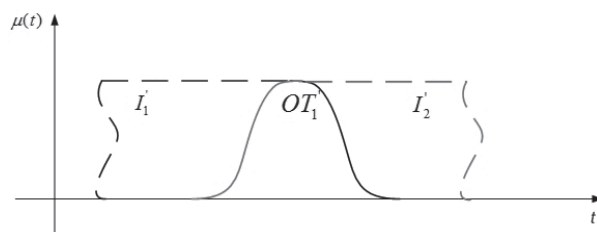


Рис. 1.1. Представление нечеткой продолжительности как пересечения нечетких интервалов

Для сравнения необработанного четкого временного ряда с нечетким сценарием необходимо восстановить нечеткие интервалы I_1', \dots, I_{N_w}' . Исходя из условия выпуклости, неубывающий участок функции принадлежности нечеткого числа OT_1' можем считать левой границей интервала I_2' , а невозрастающий – правой границей интервала I_1' . Обобщая вышеуказанное правило на все волны сценария, получим:

$$OT'_i = I'_i \cap I'_{i+1}, i \in (1, N_w - 1), \mu_{I'_j}(t) = \begin{cases} \mu_{OT'_{j-1}}(t), t < \inf M(OT'_{j-1}), \\ 1, \inf M(OT'_{j-1}) \leq t \leq \sup M(OT'_j), \\ \mu_{OT'_j}(t), \sup M(OT'_j) < t, \end{cases}$$

$$j \in [2, N_w], \mu_{I'_i}(t) = \begin{cases} 1, 0 \leq t \leq \sup M(OT'_1), \\ \mu_{OT'_1}(t), \sup M(OT'_1) < t, \end{cases}$$

где $M(OT'_j)$ – интервал модальных значений нечеткого числа OT'_j .

Аналогично нечетким интервалам времени, для каждой волны определяются нечеткие уровни сигнала в конце волны и интервалы абсолютного значения сигнала U'_i .

$$OA'_i = \sum_{j=1}^i A'_j, j \in [1, N_w], OA'_i = U'_i \cap U'_{i+1}, i \in (1, N_w - 1),$$

$$\mu_{U'_j}(x) = \begin{cases} \mu_{OA'_{j-1}}(x), x < \min[\inf M(OA'_{j-1}), \inf M(OA'_j)], \\ 1, \min[\inf M(OA'_{j-1}), \inf M(OA'_j)] \leq x \leq \max[\sup M(OA'_{j-1}), \sup M(OA'_j)], \\ \mu_{OA'_j}(x), \max[\sup M(OA'_{j-1}), \sup M(OA'_j)] < x, \end{cases}$$

$$j \in [2, N_w], \mu_{U'_i}(x) = \begin{cases} 1, 0 \leq x \leq \sup M(OA'_1), \\ \mu_{OA'_1}(x), \sup M(OA'_1) < x. \end{cases}$$

Пусть теперь имеем некоторый фрагмент временного ряда X в виде набора отсчетов случайной величины в последовательные моменты времени:

$$X = \langle X_i \rangle = \langle (x(t_i) = x_i, t_i) \rangle, i \in [1, n], x_i \in \mathfrak{R}, t_i \in \mathfrak{R}^+,$$

$$\forall i, j \in [1, n], i < j: t_i < t_j, t_1 = t_{\min}, t_n = t_{\max}.$$

Степень соответствия некоторого отсчета (x_i, t_i) фрагмента временного ряда волне W_j ($j \in [1, N_w]$) означенного сценария обозначим как $m(W_j, X_i)$. Будем ее вычислять исходя из требования одновременной принадлежности уровня сигнала интервалу абсолютного значения сигнала волны и момента времени интервалу времени волны. При этом сам характер изменений сигнала должен соответствовать заявленной монотонной (с заданным уровнем чувствительности) зависимости от времени. Будем требовать примерного соответствия прироста сигнала относительно начала волны длине временного отрезка с начала волны:

$$\frac{x_i - x_1 - OA'_{j-1}}{t_i - t_{\min} - OT'_{j-1}} \approx \frac{A'_j}{T'_j}, j \in [2, N_w], i \in [2, n] \quad (2.1)$$

$$\frac{x_i - x_1}{t_i - t_{\min}} \approx \frac{A'_j}{T'_j}, j = 1, i \in [2, n]$$

Символ примерного равенства в выражениях (2.1) означает вычисление показателя близости. Пример такого показателя предложен в [Зернов, 2010]:

$$m_{\text{coen}}(\tilde{X}, \tilde{Y}) = \sup_{x \in \mathbb{R}} \min(\mu_x(x), \mu_y(x)). \quad (2.2)$$

Из (2.1) и (2.2) получим степень соответствия отсчета волне сценария:

$$m(W_j, X_i) = \wedge \left[\begin{array}{l} \mu_{i_j}(t_i - t_{\min}), \mu_{U_j}(x_i - x_{\min}), \\ m_{\text{coen}} \left(\frac{x_i - x_1 - OA'_{j-1}}{t_i - t_{\min} - OT'_{j-1}}, \frac{A'_j}{T'_j} \right) \end{array} \right], j \in [2, N_w], i \in [2, n] \quad (2.3)$$

$$m(W_j, X_i) = \wedge \left[\begin{array}{l} \mu_{i_j}(t_i - t_{\min}), \mu_{U_j}(x_i - x_{\min}), \\ m_{\text{coen}} \left(\frac{x_i - x_1}{t_i - t_{\min}}, \frac{A'_j}{T'_j} \right) \end{array} \right], j = 1, i \in [2, n] \quad (2.4)$$

здесь \wedge – операция Т-нормы, для X_i , $m(W_j, X_i) = 1$ по умолчанию.

Выражения (2.3)–(2.4) для степени соответствия отсчета ряда волне сценария могут давать существенную ошибку идентификации при случайных возмущениях сигнала в области начала волны. Чтобы этого избежать, можно предложить вариант расчета степени близости на основе сравнения не относительных приращений, а текущего и ожидаемого значений:

$$m(W_j, X_i) = \wedge \left[\mu_{i_j}(t_i - t_{\min}), \mu_{U_j}(x_i - x_{\min}), \mu_{D_{i,j}}(x_i - x_1) \right],$$

$$D_{i,j} = \begin{cases} OA'_{j-1} + (t_i - t_{\min} - OT'_{j-1}) \frac{A'_j}{T'_j}, j \in [2, N_w], i \in [2, n], \\ (t_i - t_{\min}) \frac{A'_j}{T'_j}, j = 1, i \in [2, n], \end{cases}$$

Сформулируем теперь степень соответствия рассматриваемого фрагмента временного ряда X сценарию P в целом – $R(P, X)$, исходя из требования, чтобы номера волн, соответствующие последовательным отсчетам фрагмента ряда, образовали неубывающую последовательность без пропусков (при условии, что отсчеты взяты с достаточной дискретностью по времени).

Значения $m(W_j, X_i)$ фактически представляют собой принадлежности волн W_j нечеткому множеству волн $WX(X_i)$, подходящих для описания X_i :

$$WX(X_i) = \{m(W_j, X_i) / W_j\}, i \in [1, n], j = \overline{1, N_w}$$

Тогда α -уровень рассмотренного множества определяется так:

$$[WX(X_i)]_\alpha = \{W_j \mid m(W_j, X_i) \geq \alpha\}, i \in [1, n], j = \overline{1, N_w}$$

Тогда $R(P, X)$ можно определить следующим образом:

$$R(P, X) = \sup_{\alpha \in [0,1]} \left\{ \alpha \left| \begin{array}{l} \exists (W_{h_1}, W_{h_2}, \dots, W_{h_n}): W_{h_i} \in [WX_i]_\alpha, i \in [1, n] \\ h_1 = 1, h_j \leq h_{j+1}, h_{j+1} - h_j \in \{0, 1\}, j \in [1, n-1] \end{array} \right. \right\}$$

При этом далеко не все волны сценария (начиная с конца) могут быть представлены среди W_{h_i} . Фактически показатель $R(P, X)$ говорит о степени соответствия рассматриваемого фрагмента временного ряда начальной части нечеткого сценария. На практике его полезно дополнить номером $H(P, X)$ волны (или множеством номеров), крайней в последовательности, обеспечившей полученное значение $R(P, X)$.

2.2. Способ идентификации означенного нечеткого сценария относительно обработанного фрагмента четкого временного ряда

Рассмотрим теперь случай, когда фрагмент временного ряда X предварительно фильтруется с целью выделить монотонные участки или участки с примерно постоянной производной.

Обработанный временной ряд тогда можно представить как означенный сценарий, составленный из четких волн:

$$PX = (WX_1, WX_2, \dots, WX_{N_{wx}}), WX_i = \langle \{1/a_i\}, \{1/t_i\}, 0 \rangle, i = \overline{1, N_{wx}}$$

Сценарий P получает означивание только в том случае, если в сценарии PX присутствует волна с номером, превышающим номер основной волны из P (волну с номером N_{wx} считаем незавершенной):

$$N_{wx} > \max \{i, i \in [1, N_w], W_i \in B(P)\}.$$

Результат означивания тогда можно представить в виде:

$$P(W_{b_1} := (\{1/a_{b_1}\}, \{1/t_{b_1}\}), \dots, W_{b_{|B|}} := (\{1/a_{b_i}\}, \{1/t_{b_i}\})) W_{b_i} \in B, i \in [1, \|B\|].$$

В результате четкий сценарий PX соответствует нечеткому сценарию P , если каждая его волна, кроме последней, соответствует волне с тем же номером нечеткого сценария по величине амплитуды и продолжительности. Последняя волна при этом соответствует волне нечеткого сценария с тем же номером в смысле попадания амплитуды и продолжительности незавершенной волны в нечеткие интервалы от нуля до абсолютной амплитуды и абсолютной продолжительности волны с номером N_{wx} :

$$a_{N_{wx}} \in IA'_{N_{wx}}, \quad t_{N_{wx}} \in IT'_{N_{wx}},$$

$$\mu_{IA'_{N_{wx}}}(x) = \begin{cases} 1, 0 \leq x \leq \sup M(A'_{N_{wx}}), \\ \mu_{A'_{N_{wx}}}(x), \sup M(A'_{N_{wx}}) < x, \end{cases} \quad \mu_{IT'_{N_{wx}}}(t) = \begin{cases} 1, 0 \leq t \leq \sup M(T'_{N_{wx}}), \\ \mu_{T'_{N_{wx}}}(t), \sup M(T'_{N_{wx}}) < t. \end{cases}$$

Показатель $R(P, PX)$ тогда определяется как

$$R(P, PX) = (\mu_{IA'_{N_{wx}}}(x) \wedge \mu_{IT'_{N_{wx}}}(t)) \wedge \bigwedge_{i=1}^{N_{wx}-1} (\mu_{A_i}(x) \wedge \mu_{T_i}(t)).$$

$H(P, X)$ в данном случае всегда тривиально и составляет $\{N_{wx}\}$.

Заключение

Предложенная модель нечеткого сценария развития временного ряда позволяет описать нечеткий сценарий инвариантно относительно масштаба как по времени, так и по амплитуде сигнала.

Разработаны способы применения предложенной модели нечеткого временного ряда при идентификации нечетких сценариев для случаев означенного нечеткого сценария и фрагмента четкого временного ряда – совокупности отсчетов сигнала, а также означенно-

го нечеткого сценария и фрагмента четкого временного ряда – совокупности четких волн.

СПИСОК ЛИТЕРАТУРЫ

[Зернов, 2010] Зернов М.М. Оценка состояния объекта управления в случае нечетко-множественной входной информации //Математическая морфология. Электронный математический и медико-биологический журнал. – Т. 6. – Вып. 2. – 2007. Режим доступа к журн.: <http://www.smolensk.ru/user/sgma/MMORPH/TITL.HTM>.

[Зернов, 2014] Зернов М.М., Чуль А.Д. Модель нечеткого прецедента для идентификации реализаций паттерна волновой теории Эллиотта// Известия СмолГУ. – Смоленск: Издательство СмолГУ, 2014. №2 (26).

[Ковалев, 2007] Ковалев С.М. Гибридные нечетко-темпоральные модели временных рядов в задачах анализа и идентификации слабо формализованных процессов // Интегрированные модели и мягкие вычисления в искусственном интеллекте: сборник трудов IV Международной научно-практической конференции. В 2-х томах. Т.1 – М. : Физматлит, 2007.

[Ярушкина и др., 2010] Ярушкина Н.Г., Афанасьева Т.В., Перфильева И.Г. Интеллектуальный анализ временных рядов: учебное пособие. – Ульяновск: УлГТУ, 2010.

УДК 007:519.816

НЕЧЕТКО-СТОХАСТИЧЕСКИЙ ВЫВОД В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ ОБРАБОТКИ ПЕРВИЧНОЙ ИНФОРМАЦИИ¹

С.М. Ковалев (*ksm@rfnias.ru*)
РГУ ПС

В работе развивается концепция интеллектуального пропроцессинга и когнитивных измерений информации на основе использования схем нечетко-стохастического вывода в интересах создания новых классов прикладных интеллектуальных систем динамического типа. Предлагается схема нечетко-стохастического вывода для нечетких динамических систем, допускающих аналитическое представление в виде функциональных зависимостей.

Введение

Развитие в России современных средств автоматизации приводит к появлению интегрированных интеллектуальных систем, в которых измерения, первичная обработка информации и выработка решений осуществляются в рамках единой интеллектуальной среды [Aarts, 2001], располагающей, помимо средств сбора и обработки первичной информации, средствами интеллектуального анализа данных и извлечения знаний. Являясь базовыми для последних, ключевую роль в таких интеллектуальных средах играют средства интеллектуальных измерений (ИИ) [Romanov, 1994] и интеллектуального препроцессинга (ИП) [Ковалев, 2008].

Концепция ИИ и ИП опирается на процесс иерархического гранулирования информации [Тарасов, 2014], в соответствии с которым на нижнем уровне осуществляются детализированные измерения количественной информации, а на верхнем уровне полученные результаты отображаются на прагматическую шкалу укрупненных оценок, представленную относительно небольшим числом лингвистических значений, достаточных для принятия решений с допустимым уровнем неточности. При формировании такой

¹ Работа выполнена при финансовой поддержке РФФИ (проекты №№ 13-07-00183-а, 13-07-13108-офи-м РЖД, 13-08-12151-офи-м).

шкалы и разработке методов отображения на лингвистическую шкалу мелкозернистой информации следует учитывать две важные особенности первичной информации. Первая связана с динамической природой первичной информации, обусловленной развитием во времени контролируемых или управляемых технологических процессов, инициирующих появление первичных данных на выходе измерительных систем. Динамическая природа информации требует при реализации ИИ и ИП осуществлять упреждающее прогнозирование данных на выходе измерительных систем с учетом имеющихся знаний о процессах порождения первичной информации. Вторая особенность связана с наличием фактора неопределенности как неотъемлемого атрибута любых измерений. Он включает в себя множество различных, так называемых Не-факторов знаний [Нариньяни, 1994], обусловленных стохастической природой информации, неполнотой и неточностью знаний о свойствах объекта и условиях проведения измерений, нечеткостью критериев достоверности измерений. Для интеллектуальных измерений в технических и организационно-технологических системах наиболее характерными НЕ-факторами являются стохастическая неопределенность, обусловленная несовершенством приборов и влиянием шумовых процессов в измерительной аппаратуре, и нечеткие факторы, обусловленные неполнотой знаний о процессе порождения первичной информации на выходе измерительной системы и нечетко-субъективной природой приемника информации – лица, принимающего решения.

С учетом сказанного одной из ключевых задач в рамках разработки общей методологии ИИ и ИП как иерархического процесса грануляции первичных данных является гранулирование динамической информации с учетом упомянутых выше атрибутов неопределенности измерений. В настоящей работе предлагается новый подход к гранулированию динамической информации, основанный на использовании специальной схемы нечетко-стохастического вывода.

1. Схема нечетко-стохастического вывода

Объектом ИИ и ИП является первичная информация, появляющаяся на выходе сенсорных и измерительных систем в виде временного ряда наблюдаемых числовых значений $X = (x_1, x_2, \dots, x_i, \dots)$.

Будем предполагать, что эти данные являются внешним проявлением развития некоторого динамического процесса, представленного разностным уравнением

$$x_{i+1} = f(x_i, \varepsilon, C), \quad (1)$$

где $f(x_i, \varepsilon, C)$ – нелинейная функция (отображение), ε – шум измерений с известной плотностью распределения вероятности $p_\varepsilon(\varepsilon)$, C – вектор параметров.

Рассмотрим механизм влияния двух основных факторов неопределенности – нечеткого и стохастического – на информацию, содержащуюся в стохастическом процессе, порожденном отображением (1), и способ описания взаимодействия этих факторов, положенный в основу нечетко-стохастического вывода.

Стохастическая неопределенность обусловлена воздействием на наблюдаемые значения x_i шумового процесса, модель которого представлена случайной величиной $x_i = x(t_i)$ с плотностью распределения вероятностей $p_\varepsilon(\varepsilon)$. Воздействие шума на наблюдаемые значения x_i представляется сложением детерминированной величины $X = x_i$ и случайной величины ε , результатом которого является случайная величина с плотностью распределения

$$\rho(x) = \rho(x - x_i). \quad (2)$$

Случайные числовые значения x поступают на вход приемника информации, в качестве которого выступает ЛППР, обладающий знаниями о процессе порождения наблюдаемой на выходе измерительной системы информации и минимально-необходимом уровне ее детализации с учетом конечной цели принятия решений. Например, опытный электромеханик, обслуживающий исполнительные устройства в системах горочной автоматизации на железнодорожных сортировочных станциях, имеет достаточно полное представление о динамике изменения давления в пневмосети вагонного замедлителя (ВГЗ – устройства торможения скатывающихся с горки отцепов). При этом в данных, поступающих с электропневматических датчиков, ключевой информацией для принятия решений (диагностирование ВГЗ, смена режима управления и др.) является вероятность появления событий, связанных с резкими изменениями давления в пневмосети ВГЗ, то есть с

переходом «больших» значений давления на «малые» и наоборот. Поэтому в данном конкретном случае гранулирование информации осуществляется по атрибуту нечеткости с точностью до «больших» и «малых» значений измеряемого давления, а по стохастическому фактору – с точностью до высокой и низкой вероятности появления нечеткого события, связанного с «резким скачком» давления в пневмосети ВГЗ.

Модель интеллектуального приемника первичной информации, обладающего знаниями о процессе порождения информации, заявим в виде нечетко-динамической системы (НДС), состоящей из множества нечетких продукционных правил. Левые части правил содержат условия, накладываемые на значения переменных состояния процесса x_i в i -й момент времени, а правые части – нечеткие гранулированные оценки этих значений в прогнозируемый последующий $(i+1)$ -ый момент времени. Совместное влияние нечеткостного и стохастического факторов рассмотрим на примере нечеткой динамической системы типа Такаги-Сугено [Takagi, 1985], представленной множеством нечетких продукционных правил:

$$\begin{aligned} \text{ЕСЛИ } x_i = \alpha_1 \text{ ТО } x_{i+1} &= c_{01} + c_{11}x_i + c_{21} \cdot x_i^2 + \dots + c_{n1}x^n \\ \text{ЕСЛИ } x_i = \alpha_2 \text{ ТО } x_{i+1} &= c_{02} + c_{12}x_i + c_{22} \cdot x_i^2 + \dots + c_{n2}x^n \\ &\dots \\ \text{ЕСЛИ } x_i = \alpha_n \text{ ТО } x_{i+1} &= c_{0n} + c_{1n}x_i + c_{2n} \cdot x_i^2 + \dots + c_{nn}x^n, \end{aligned}$$

где x_i, x_{i+1} – наблюдаемые значения переменной состояния процесса в текущий i -й и последующий $(i+1)$ -й моменты времени; α_i – лингвистические значения (нечеткие термы) переменной состояния процесса, определенные на числовой шкале X при помощи соответствующих функций принадлежности (ФП).

Пусть на вход НДС поступает значение случайной величины x_i с плотностью распределения $\rho(x)$, вычисляемой на основании (2). Поскольку нечеткая система Сугено допускает аналитическое представление в виде функциональной зависимости $x_{i+1} = f(x_i, C)$ (C – вектор параметров НДС), стохастическое распределение $\rho(x)$, действующее на входе НДС, индуцирует на выходе НДС стохастическое распределение $\gamma(x)$, которое можно определить на основе закона суперпозиции функций распределения. Аналитическим представлением НДС является функциональная зависимость:

$$x_{i+1} = f(x_i, C) = \frac{\sum_{k=1}^n \mu_{\alpha k}(x_i) \cdot (c_{0k} + c_{1k} \cdot x_i + \dots + c_{nk} \cdot x_i^n)}{\sum_{k=1}^n \mu_{\alpha k}(x_i)}, \quad (3)$$

где $C = (c_{01}, c_{11}, \dots, c_{n1}, \dots, c_{m1})$ – вектор параметров НДС.

Если предположить монотонность функции $f(x_i, C)$, стохастическое распределение величины x_{i+1} на выходе НДС определяется на основе закона суперпозиции функций распределения как [Вентцель, 2005]:

$$\gamma(x_{i+1}) = \rho(f^{-1}(x_{i+1}, C)) \cdot \left| \frac{df^{-1}(x_{i+1}, C)}{dx_{i+1}} \right|, \quad (4)$$

где $f^{-1}(x_{i+1}, C)$ – функция, обратная функции $f(x_i, C)$.

Для немонотонной функции $f(x_i, C)$ стохастическое распределение выходной величины определяется как сумма частных распределений по всем монотонным участкам:

$$\gamma(x_{i+1}) = \sum_l \rho(f_l^{-1}(x_{i+1}, C)) \cdot \left| \frac{df_l^{-1}(x_{i+1}, C)}{dx_{i+1}} \right|, \quad (5)$$

где $f_l^{-1}(x_{i+1}, C)$ – часть функции $f^{-1}(x_{i+1}, C)$, определенная на l -ом участке ее монотонности.

Пусть x_i – текущее случайное значение переменной состояния процесса в i -й момент времени с плотностью распределения $\rho(x)$, x_{i+1} – прогнозируемое состояние процесса в $(i+1)$ -й момент времени с плотностью распределения $\gamma(x)$, α и β – нечеткие термы (нечеткие числа), заданные на числовой шкале X в виде ФП $\mu_\alpha(x)$ и $\mu_\beta(x)$. Вычислим вероятности появления нечетких событий $x_i = \alpha$ и $x_{i+1} = \beta$, связанных с наблюдением нечетких значений α и β переменной состояния процесса X в моменты времени i и $i+1$ соответственно. Введем в рассмотрение величину $\varepsilon \in [0, 1]$ в качестве некоторой оценки порога истинности для вышеупомянутых нечетких событий. Вероятности появления нечетких событий определим путем гранулирования числовых значений переменной состояния X с использованием ε -уровней ФП $\mu_\alpha(x)$ и $\mu_\beta(x)$ и последующего вычисления вероятностей попадания значений x_i и x_{i+1} в интервалы соответствующих ε -уровней. Для заданного порога истинности ε нечеткому терму α на шкале X соответствует ε -уровень в виде

интервала $\Delta_\alpha \subset X$ такого, что $\forall x \in \Delta_\alpha \mu_\alpha(x) \geq \varepsilon$. В пределах интервала Δ_α все значения $x \in \Delta_\alpha$ считаются равными нечеткому числу α с заданной степенью истинности ε . Вероятность появления нечеткого события $x_i = \alpha$, связанного с наблюдением нечеткого числового значения α , равна вероятности принадлежности значения x_i интервалу Δ_α :

$$P(x_i = \alpha) = \int_{x \in \Delta_\alpha} \rho(x) dx, \quad (6)$$

где $\rho(x)$ – плотность распределения вероятностей наблюдаемого текущего значения x_i .

Аналогично определяется вероятность появления нечеткого события $x_{i+1} = \beta$ в последующий момент времени $i+1$. Определяется ε -уровень ФП $\mu_\beta(x)$ нечеткого термина β в виде интервала Δ_β . Вероятность нечеткого события $P(x_{i+1} = \beta)$ равна вероятности принадлежности значения x_{i+1} интервалу Δ_β :

$$P(x_{i+1} = \beta) = \int_{x \in \Delta_\beta} \gamma(x) dx, \quad (7)$$

где $\gamma(x)$ – плотность распределения вероятностей прогнозируемого значения x_{i+1} .

Формулы (4) и (5), описывающие процесс трансформации стохастического распределения $\rho(x)$ случайной величины x_i на входе НДС в нечетко-стохастическое распределение $\gamma(x)$ выходной величины x_{i+1} , определяют схему стохастического вывода

$$\frac{\rho(x_i), x_i \rightarrow x_{i+1}}{\gamma(x_{i+1})} \quad (8)$$

Схема (8) вместе с формулами (6) и (7) образует схему нечетко-стохастического вывода:

$$\frac{P(x_i = \alpha), x_i \rightarrow x_{i+1}}{P(x_{i+1} = \beta)} \quad (9)$$

Следует отметить, что вышеприведенные схемы нечетко-стохастического вывода справедливы для НДС, допускающих аналитическое представление в виде функциональной зависимости. В более общем случае, когда выходные значения нечеткой системы представлены нечеткими множествами, например для нечетких систем типа Мамдани [7], результатом нечетко-стохастического вывода являются двумерные распределения.

2. Поясняющий пример

Процесс порождения первичной информации представлен в виде простейшей нечетко-динамической системы

$$\begin{aligned} \text{Если } x_i = \text{ОКОЛО НУЛЯ, ТО } x_{i+1} &= 4 \\ \text{Если } x_i = \text{ОКОЛО ТРЕХ, ТО } x_{i+1} &= 10 \end{aligned}$$

При линейных ФП нечетких термов *ОКОЛО НУЛЯ* и *ОКОЛО ТРЕХ* аналитическим представлением НДС является линейное разностное уравнение

$$x_{i+1} = 2x_i + 4. \quad (10)$$

Наблюдаемые состояния процесса подвержены влиянию шума с треугольной функцией плотности распределения вероятностей $\rho(x)$, нулевым математическим ожиданием и дисперсией, равной 2. В i -й момент времени наблюдаемое состояние процесса $x_i = 2$. Требуется определить вероятность «скачка» – вероятность перехода МАЛОГО значения состояния процесса, наблюдаемого в i -й момент времени, в БОЛЬШОЕ значение, прогнозируемое к последующему $i+1$ -му моменту времени.

В i -й момент времени на входе НДС действует наблюдаемое состояние процесса с функцией плотности распределения $\rho(x)$:

$$\rho(x) = \begin{cases} x_i - 1 & \text{при } 1 < x \leq 2 \\ -x + 3 & \text{при } 2 < x \leq 3 \\ 0 & \text{при } x \leq 1 \text{ либо } x > 3 \end{cases}.$$

На отрезке $[1,3]$ функция (10), описывающая отображение ВХОД – ВЫХОД НДС, монотонная, поэтому плотность распределения прогнозируемого значения на выходе НДС определяется на основании (4):

$$\gamma(x) = \rho(f^{-1}(x)) \cdot \left| \frac{df^{-1}(x)}{dx} \right|,$$

$$\text{где } f^{-1}(x) = \frac{x-4}{2}, \quad \left| \frac{df^{-1}(x)}{dx} \right| = \frac{1}{2}.$$

В результате получаем функцию плотности распределения вероятности прогнозируемого значения:

$$\gamma(x) = \begin{cases} \frac{x}{4} - \frac{3}{2} & \text{при } 6 < x \leq 8 \\ -\frac{x}{4} + \frac{5}{2} & \text{при } 8 < x \leq 10 \\ 0 & \text{при } x \leq 6 \text{ либо } x > 10 \end{cases} .$$

Вычислим вероятности появления нечетких событий α ; " $x_i = \text{МАЛОЕ}$ " и β : " $x_{i+1} = \text{БОЛЬШОЕ}$ " для заданного порога истинности $\varepsilon = 0,7$ и треугольных ФП нечетких термов α и β , представленных на интервале $[0, 10]$ в виде уравнений

$$\mu_{\alpha}(x_i) = 0.1 \cdot x_i \quad \mu_{\beta}(x_{i+1}) = 1 - 0.1 \cdot x_i$$

Для порога истинности $\varepsilon = 0,7$ нечетким термам α и β на шкале X соответствуют два ε -уровня: $\Delta_{\alpha} = [0,3]$ и $\Delta_{\beta} = [7,10]$. Вероятности появления нечетких событий α ; " $x_i = \text{МАЛОЕ}$ " и β : " $x_{i+1} = \text{БОЛЬШОЕ}$ " вычисляются на основании формул (6) и (7):

$$P(x_i = \text{МАЛОЕ}) = \int_1^2 (x_i - 1) dx_i + \int_2^3 (3 - x_i) dx_i = 1,$$

$$P(x_{i+1} = \text{БОЛЬШОЕ}) = \int_6^8 \left(\frac{x_{i+1}}{4} - \frac{3}{2}\right) dx_{i+1} + \int_8^{10} \left(\frac{5}{2} - \frac{x_{i+1}}{4}\right) dx_{i+1} \approx 0,88.$$

Вероятность перехода МАЛОГО значения состояния процесса x_i в БОЛЬШОЕ значение x_{i+1} можно выразить через произведение вероятностей отдельных событий $P(x_i = \alpha) \cdot P(x_{i+1} = \beta) = 0,88$.

Заключение

В работе развивается концепция интеллектуального пропроцессинга информации на основе использования схем нечетко-стохастического вывода. Предлагается схема нечетко-стохастического вывода для нечетких динамических систем, допускающих аналитическое представление в виде функциональных зависимостей. Дальнейшие перспективы исследования связаны с обобщением нечетко-стохастической схемы вывода на другие типы нечетких систем и их применением для извлечения знаний из потоковых и темпоральных данных.

СПИСОК ЛИТЕРАТУРЫ

- [**Aarts, 2001**] Aarts E., Harwig, R., Schuurmans, M. Ambient Intelligence// The Invisible Future: The Seamless Integration of Technology into Everyday Life/ Ed. by P.J.Denning. – New York: McGraw-Hill Companies, 2001.
- [**Romanov, 1994**] Romanov, V.N., Sobolev, V.S., Tsvetkov, E.I. Intelligent Measurement Tools/ Ed. by E.I. Tsvetkov – Moscow: Tatyani den', 1994.
- [**Ковалев, 2008**] С.М. Ковалев, А.Е. Хатламаджиян, А.И. Долгий. Интеллектуальный препроцессинг георадиолокационных данных на основе нейросетевой модели поиска паттернов временных рядов // Известия ВГТУ. Актуальные проблемы управления вычислительной техники и информатики в вычислительных системах. Вып. 5, №8 (46), 2008. С. 65–70.
- [**Тарасов, 2014**] В.Б. Тарасов, М.Н. Святкина. Когнитивные измерения в интеллектуальных системах мониторинга объектов железнодорожной инфраструктуры // Вестник РГУПС, №1, 2014. С. 72–79.
- [**Нариньяни, 1994**] Нариньяни А.С. НЕ-факторы и инженерия знаний: от наивной формализации к естественной прагматике// Сборник трудов IV Национальной конференции по ИИ (КИИ-94, Рыбинск, сентябрь 1994 г.). Т.1. – Тверь: АИИ, 1994. – С. 9–18.
- [**Takagi, 1985**] Takagi T., Sugeno M. Fuzzy identification of systems and its application to modeling and control, IEEE Trans. Syst, Man, Cybern. 1985. vol. 15. 116–132.
- [**Mamdani, 1974**] Mamdani E.H. Applications of fuzzy algorithms for control of simple dynamic plant. Proceedings IEEE. 1974, № 121 (12), pp. 1585-1588.
- [**Вентцель, 2005**] Вентцель Е.С. Теория вероятностей: Учебник для студ. вузов. – М.: Издательский центр «Академия», 2005. – 576 с.

УДК 004.8

МЕТОД ВЫВОДА НА ОСНОВЕ НЕЧЕТКОЙ СТЕПЕНИ ИСТИННОСТИ ДЛЯ СИСТЕМ СО МНОГИМИ НЕЧЕТКИМИ ВХОДАМИ¹

Д.А. Куценко (*dakutsenko@mail.ru*)

В.М. Поляков (*p_v_m@mail.ru*)

В.Г. Синюк (*vgsinuk@mail.ru*)

БГТУ им. В.Г. Шухова, Белгород

В работе описывается метод вывода на основе нечеткой степени истинности для нечетких продукционных моделей систем со многими нечеткими входами. Проводится сравнение метода с исходным методом Заде и популярным методом Мамдани. Доказывается вычислительная эффективность предложенного метода.

Введение

За последние десятилетия было предложено множество различных методов нечеткого логического вывода. Согласно классификации, приведенной в [Рутковский, 2010], все методы можно разделить на три типа – логического, типа Мамдани и типа Такаги–Сугено. Второй тип связан с работой Э. Мамдани [Mamdani, 1974], в которой был предложен эффективный метод вывода для случая, когда на входы системы поступают четкие скалярные значения. В последнее время неоднократно отмечалось (см., например, [Dubois et al., 2012]), что метод Мамдани может оказаться неадекватным для решения задач, когда входные значения представляют собой нечеткие множества. Такие задачи возникают в случае, когда исходные данные являются нечеткими либо по своей природе, либо обладают другими НЕ-факторами [Душкин, 2011], такими как неточность, неопределенность и недоопределенность, которые могут быть преобразованы в нечеткость.

В данной статье рассматриваются методы вывода с полиномиальной вычислительной сложностью для нечетких систем логического типа и типа Мамдани при n нечетких входах.

¹ Работа выполнена при финансовой поддержке РФФИ (гранты № 14-07-00154-А и № 12-07-00493-А).

1. Постановка задачи

Задача, которая решается с помощью нечеткой продукционной системы, формулируется следующим образом. Рассмотрим систему с n входами x_1, \dots, x_n и одним выходом y , представленными одноименными лингвистическими переменными. Пусть X_i – базовое множество значений i -го входа системы, $x_i \in X_i, i = \overline{1, n}$; Y – базовое множество значений выхода системы, $y \in Y$. Введем следующие обозначения. Пусть $\text{Fuzzy}(Z)$ – множество, состоящее из всех нечетких подмножеств множества Z , $\mu_{\tilde{F}}: Z \rightarrow [0, 1]$ – функция принадлежности нечеткого множества $\tilde{F} \in \text{Fuzzy}(Z)$. Под нечетким значением истинности будем понимать нечеткое подмножество $[0, 1]$, где отрезок $[0, 1]$ представляет собой множество истинностных значений, как это принято в многозначных логиках. Для нечеткого значения истинности $\tilde{V} \in \text{Fuzzy}([0, 1])$ введем обозначение $\tau_{\tilde{V}}: [0, 1] \rightarrow [0, 1]$. Взаимосвязь входов и выхода в системе описывается с помощью N нечетких правил следующего вида:

$$\tilde{R}_k: \text{Если } \langle x_1 \text{ есть } \tilde{A}_{1k} \rangle \text{ и } \dots \text{ и } \langle x_n \text{ есть } \tilde{A}_{nk} \rangle, \text{ то } \langle y \text{ есть } \tilde{B}_k \rangle, k = \overline{1, N}, \quad (1.1)$$

где $\tilde{A}_{1k} \in \text{Fuzzy}(X_1), \dots, \tilde{A}_{nk} \in \text{Fuzzy}(X_n)$ – термы из терм-множеств входных лингвистических переменных x_1, \dots, x_n соответственно, из которых сформирован антецедент k -го правила; $\tilde{B}_k \in \text{Fuzzy}(Y)$ – терм из терм-множества выходной лингвистической переменной y , с помощью которого образуется консеквент k -го правила. Пусть заданы нечеткие множества $\tilde{A}'_i \in \text{Fuzzy}(X_i), i = \overline{1, n}$, представляющие собой реальные значения входов x_1, \dots, x_n . Задача заключается в определении выхода системы $\tilde{B}' \in \text{Fuzzy}(Y)$.

Особенность систем логического типа, согласно классификации, приведенной в [Рутковский, 2010], заключается в том, что правило (1.1) формализуется с использованием нечеткой импликации в виде нечеткого отношения $\tilde{R}_k \in \text{Fuzzy}(X_1 \times \dots \times X_n \times Y)$ следующим образом:

$$\tilde{R}_k = (\tilde{A}_{1k} \times \dots \times \tilde{A}_{nk}) \Rightarrow \tilde{B}_k, \quad k = \overline{1, N},$$

где $\langle \Rightarrow \rangle$ – нечеткая импликация, выражающая причинно-следственную связь между антецедентом $\langle \langle x_1 \text{ есть } \tilde{A}_{1k} \rangle \text{ и } \dots \text{ и } \langle x_n \text{ есть } \tilde{A}_{nk} \rangle \rangle$ и консеквентом $\langle \langle y \text{ есть } \tilde{B}_k \rangle \rangle$ k -го правила.

нечеткие значения $\tilde{A}'_1, \dots, \tilde{A}'_n$ соответственно, вывод выходного нечеткого значения \tilde{B}' принимает вид

$$\begin{aligned} \text{Если } \langle x_1 \text{ есть } \tilde{A}'_1 \rangle \text{ и } \dots \text{ и } \langle x_n \text{ есть } \tilde{A}'_n \rangle, \text{ то } \langle y \text{ есть } \tilde{B}' \rangle, \\ \langle x_1 \text{ есть } \tilde{A}'_1 \rangle \text{ и } \dots \text{ и } \langle x_n \text{ есть } \tilde{A}'_n \rangle. \\ \langle y \text{ есть } \tilde{B}' \rangle. \end{aligned}$$

и осуществляется следующим образом:

$$\mu_{\tilde{B}'}(y) = \sup_{(x_1, \dots, x_n)} \left\{ \prod_{i=1, n} \{\mu_{\tilde{A}'_i}(x_i)\} * \mathbb{I} \left(\prod_{i=1, n} \{\mu_{\tilde{A}'_i}(x_i)\}, \mu_{\tilde{B}'}(y) \right) \right\}, \quad (2.3)$$

где супремум нужно брать по всем n -кам $(x_1, \dots, x_n) \in X_1 \times \dots \times X_n$, что определяет порядок сложности вычисления $O(|X|^n)$.

3. Метод вывода на основе нечеткой степени истинности

Рассмотрим соотношение (2.2). Используя правило истинностной модификации [Zadeh, 1978], [Борисов и др., 1982], [Куценко и др., 2008a], можно записать $\mu_{\tilde{A}'}(x) = \tau_{\tilde{A}|\tilde{A}'}(\mu_{\tilde{A}}(x))$, где $\tau_{\tilde{A}|\tilde{A}'}$ – нечеткая степень истинности нечеткого множества \tilde{A} относительно \tilde{A}' , представляющая собой совместимость $\text{CP}(\tilde{A}, \tilde{A}')$ терма \tilde{A} по отношению ко входному значению \tilde{A}' : $\tau_{\tilde{A}|\tilde{A}'}(\sigma) = \mu_{\text{CP}(\tilde{A}, \tilde{A}')}(\sigma) = \sup_{\substack{x \in X \\ \mu_{\tilde{A}}(x) = \sigma}} \{\mu_{\tilde{A}'}(x)\}$.

Здесь $\sigma \in [0, 1]$. Отметим, что при кусочно-линейных $\mu_{\tilde{A}}(x)$ и $\mu_{\tilde{A}'}(x)$ функцию $\tau_{\tilde{A}|\tilde{A}'}(\sigma)$ можно получить в аналитической форме с помощью метода, предложенного в [Куценко и др., 2008b].

Перейдем от переменной x к переменной σ , обозначив $\sigma = \mu_{\tilde{A}}(x)$. Получим $\mu_{\tilde{A}'}(x) = \tau_{\tilde{A}|\tilde{A}'}(\mu_{\tilde{A}}(x)) = \tau_{\tilde{A}|\tilde{A}'}(\sigma)$. Функцию принадлежности бинарного нечеткого отношения $\tilde{R} \in \text{Fuzzy}(X \times Y)$ тогда можно представить как $\mu_{\tilde{R}}(x, y) = \mathbb{I}(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)) = \mathbb{I}(\sigma, \mu_{\tilde{B}}(y))$, где \mathbb{I} – операция нечеткой импликации. Таким образом, (2.2) выражается

$$\mu_{\tilde{B}'}(y) = \sup_{\sigma \in [0, 1]} \left\{ \tau_{\tilde{A}|\tilde{A}'}(\sigma) * \mathbb{I}(\sigma, \mu_{\tilde{B}}(y)) \right\}. \quad (3.1)$$

Для обобщения (3.1) для систем с n входами можно доказать, что если \tilde{C} и \tilde{D} – нечеткие суждения, соответственно имеющие вид

$$\tilde{C} = \langle (x_1, \dots, x_n) \text{ есть } (\tilde{A}'_1 \text{ и } \dots \text{ и } \tilde{A}'_n) \rangle; \quad \tilde{D} = \langle (x_1, \dots, x_n) \text{ есть } (\tilde{A}_1 \text{ и } \dots \text{ и } \tilde{A}_n) \rangle,$$

где $\tilde{A}'_i, \tilde{A}_i \in \text{Fuzzy}(X_i)$, $i = \overline{1, n}$, то значение истинности нечеткого суждения \tilde{D} относительно \tilde{C} записывается следующим образом:

$$\tau_{\tilde{D}|\tilde{C}} = \tilde{\text{T}} \tau_{\tilde{A}_i|\tilde{A}'_i}, \quad (3.2)$$

где $\tau_{\tilde{A}_i|\tilde{A}'_i}$ – нечеткое значение истинности \tilde{A}_i относительно \tilde{A}'_i ; $\tilde{\text{T}}$ – расширенная по принципу обобщения n -местная t -норма.

Доказательство. По определению $\tau_{\tilde{D}|\tilde{C}}(\sigma) = \sup_{\substack{(x_1, \dots, x_n) \in X_1 \times \dots \times X_n \\ \mu_{\tilde{D}}(x_1, \dots, x_n) = \sigma}} \mu_{\tilde{C}}(x_1, \dots, x_n)$. Здесь $\sigma = \mu_{\tilde{D}}(x_1, \dots, x_n) = \prod_{i=1, n} \mu_{\tilde{A}_i}(x_i) = \prod_{i=1, n} \sigma_i$; $\mu_{\tilde{C}}(x_1, \dots, x_n) = \prod_{i=1, n} \mu_{\tilde{A}'_i}(x_i)$. Тогда $\tau_{\tilde{D}|\tilde{C}}(\sigma) = \sup_{\substack{(x_1, \dots, x_n) \in X_1 \times \dots \times X_n \\ \prod_{i=1, n} \sigma_i = \sigma}} \left\{ \prod_{i=1, n} \mu_{\tilde{A}'_i}(x_i) \right\}$.

Данное равенство не изменится, если его записать

$$\tau_{\tilde{D}|\tilde{C}}(\sigma) = \sup_{\substack{(x_1, \dots, x_n) \in X_1 \times \dots \times X_n \\ \prod_{i=1, n} \sigma_i = \sigma}} \left\{ \prod_{i=1, n} \left\{ \sup_{x_i \in W'_i} \mu_{\tilde{A}'_i}(x_i) \right\} \right\}, \quad \text{где } W'_i = \{x_i \in X_i \mid \mu_{\tilde{A}_i}(x_i) = \sigma_i\},$$

$X'_i = \{x_i \in X_i \mid \mu_{\tilde{A}_i}(x_i) \geq \sigma_i\}$. Учитывая, что $\sup_{x_i \in W'_i} \mu_{\tilde{A}'_i}(x_i) = \tau_{\tilde{A}_i|\tilde{A}'_i}(\sigma_i)$ – нечеткое значение истинности \tilde{A}_i по отношению к \tilde{A}'_i , можно записать функцию принадлежности нечеткой истинности $\tau_{\tilde{D}|\tilde{C}}$ как

$$\tau_{\tilde{D}|\tilde{C}}(\sigma) = \sup_{\substack{(\sigma_1, \dots, \sigma_n) \in [0, 1]^n \\ \prod_{i=1, n} \sigma_i = \sigma}} \left\{ \prod_{i=1, n} \tau_{\tilde{A}_i|\tilde{A}'_i}(\sigma_i) \right\}. \quad \text{Последнее равенство соответству-$$

ет расширенной по принципу обобщения t -норме (3.2), что и требовалось доказать.

С учетом вышеприведенного доказательства вывод выходного нечеткого значения \tilde{B}' на основе нечеткой степени истинности примет вид

$$\mu_{\tilde{B}'}(y) = \sup_{\sigma \in [0, 1]} \left\{ \mu_{\tilde{A}_i|\tilde{A}'_i}(\sigma) * \text{I}(\sigma, \mu_{\tilde{B}}(y)) \right\}. \quad (3.3)$$

Выражение (3.3) характеризуется сложностью порядка $O(n|X|^2)$.

Таким образом, применение нечеткой степени истинности для вывода в логических системах при нечетких входных значениях позволяет снизить сложность с экспоненциальной до полиномиальной.

4. Вывод нечеткого выходного значения для блока правил

Для того чтобы определить вывод для N правил (1.1), необходимо определить, какой метод будет использоваться при решении этой задачи. Рассмотрим метод типа Мамдани [Mamdani, 1974]. В этом случае агрегирование значений вывода $\tilde{B}'_1, \dots, \tilde{B}'_N$ по каждому правилу осуществляется путем объединения $\tilde{B}' = \bigcup_{k=1}^N \tilde{B}'_k$ с функцией принадлежности \tilde{B}' , вычисляемой с помощью t -нормы, то есть

$$\mu_{\tilde{B}'}(y) = \bigcup_{k=1, N} \mu_{\tilde{B}'_k}(y). \quad (4.1)$$

Если используется логическая модель, то агрегирование выходов по каждому правилу осуществляется путем пересечения $\tilde{B}' = \bigcap_{k=1}^N \tilde{B}'_k$ с использованием t -нормы, то есть

$$\mu_{\tilde{B}'}(y) = \bigcap_{k=1, N} \mu_{\tilde{B}'_k}(y). \quad (4.2)$$

Дефаззификация осуществляет отображение нечеткого множества \tilde{B}' в четкое скалярное значение \bar{y} . При использовании метода дефаззификации по среднему центру

$$\bar{y} = \frac{\sum_{l=1}^N \bar{y}_l \cdot \mu_{\tilde{B}'}(\bar{y}_l)}{\sum_{l=1}^N \mu_{\tilde{B}'}(\bar{y}_l)}, \quad (4.3)$$

где $\bar{y}_l, l = \overline{1, N}$ – значения, где находятся центры функций принадлежности $\mu_{\tilde{B}'_k}$ [Пегат, 2009].

При рассмотрении методов типа Мамдани $I(\cdot)$ функционирует как t -норма, то есть

$$I(\sigma, \mu_{\tilde{B}'_k}(y)) = T(\sigma, \mu_{\tilde{B}'_k}(y)) = \sigma * \mu_{\tilde{B}'_k}(y). \quad (4.4)$$

Нечеткое множество \tilde{B}' , которое определяет совокупный выход блока правил с учетом (3.3), (4.1) и (4.4):

$$\mu_{\tilde{B}'}(\bar{y}_l) = \bigwedge_{k=1, \bar{N}} \left\{ \sup_{\sigma \in [0, 1]} \left\{ \tau_k(\sigma) * \Gamma(\sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) \right\} \right\}.$$

где $\tau_k(\sigma) = \mu_{\tilde{A}'_k}(\sigma)$. Следовательно, отношение (4.3) принимает вид:

$$\bar{y} = \frac{\sum_{l=1}^N \bar{y}_l \cdot \bigwedge_{k=1, \bar{N}} \left\{ \sup_{\sigma \in [0, 1]} \left\{ \tau_k(\sigma) * \Gamma(\sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) \right\} \right\}}{\sum_{l=1}^N \bigwedge_{k=1, \bar{N}} \left\{ \sup_{\sigma \in [0, 1]} \left\{ \tau_k(\sigma) * \Gamma(\sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) \right\} \right\}}. \quad (4.5)$$

В оригинальном методе Мамдани [Mamdani, 1974] используется частный случай t -нормы $\Gamma = \min$ и t -конормы $S = \max$. Но при применении других t -норм, как, например,

$$\Gamma(\sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) = \max\{\sigma + \mu_{\tilde{B}_k}(\bar{y}_l) - 1, 0\}, \quad (4.6)$$

алгоритм Мамдани при нечетких входах не может быть реализован с полиномиальной вычислительной сложностью.

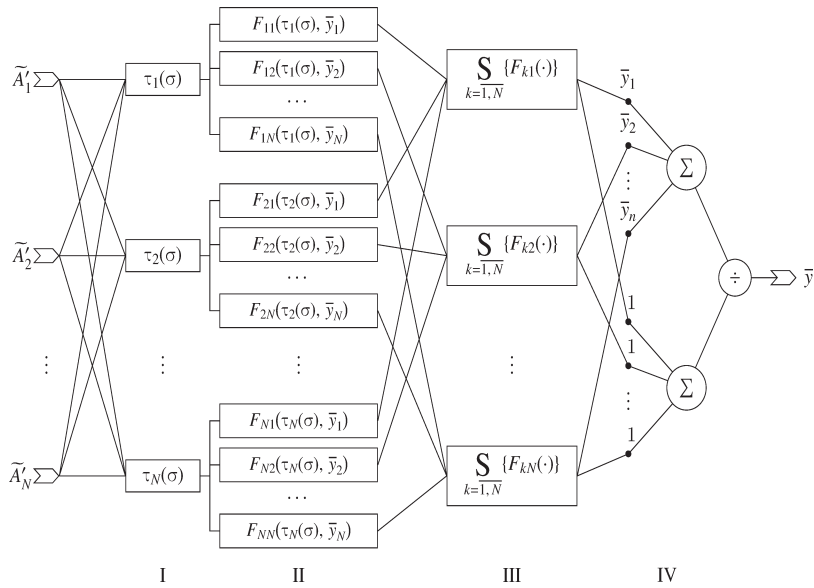


Рис. 1. Структура сети, соответствующая соотношению (4.5)

На рис. 1 соотношение (4.5) представлено в виде сетевой структуры системы, где $F_{kl}(\tau_k(\sigma), \bar{y}_l) = \sup_{\sigma \in [0,1]} \left\{ \tau_k(\sigma) * \text{T}(\sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) \right\}$.

При рассмотрении систем логического типа $I(\cdot)$ функционирует как нечёткая импликация. Далее воспользуемся S -импликацией [Waczuński et al., 2008]. В этом случае

$$I(\sigma, \mu_{\tilde{B}_k}(y)) = S(1 - \sigma, \mu_{\tilde{B}_k}(y)). \quad (4.7)$$

Следуя (3.3), (4.2) и (4.7), нечеткое множество \tilde{B}' будет определяться $\mu_{\tilde{B}'}(y) = \prod_{k=1, N} \left\{ \sup_{\sigma \in [0,1]} \left\{ \tau_k(\sigma) * S(1 - \sigma, \mu_{\tilde{B}_k}(y)) \right\} \right\}$. Отсюда выражение (4.3) принимает вид:

$$\bar{y} = \frac{\sum_{l=1}^N \bar{y}_l \cdot \prod_{k=1, N} \left\{ \sup_{\sigma \in [0,1]} \left\{ \tau_k(\sigma) * S(1 - \sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) \right\} \right\}}{\sum_{l=1}^N \prod_{k=1, N} \left\{ \sup_{\sigma \in [0,1]} \left\{ \tau_k(\sigma) * S(1 - \sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) \right\} \right\}}. \quad (4.8)$$

Структура сети, соответствующая выражению (4.8), будет совпадать со структурой на рис. 1, но поменяется содержимое II и III уровней следующим образом: $F_{kl}(\tau_k(\sigma), \bar{y}_l) = \sup_{\sigma \in [0,1]} \left\{ \tau_k(\sigma) * S(1 - \sigma, \mu_{\tilde{B}_k}(\bar{y}_l)) \right\}$;

$\bigwedge_{k=1, N} \{F_{kl}(\cdot)\}$ заменяется на $\prod_{k=1, N} \{F_{kl}(\cdot)\}$.

Заключение

Предложенный в работе метод нечеткого вывода для систем логического типа и для систем типа Мамдани, на входы которых поступают нечеткие значения, может использовать различные t -нормы для реализации операции нечеткой импликации. Метод имеет полиномиальную вычислительную сложность, что позволяет использовать его для решения задач моделирования систем с большим количеством нечетких входов.

СПИСОК ЛИТЕРАТУРЫ

- [**Борисов и др., 1982**] Борисов А.Н., Алексеев А.В., Крумберг О.А. и др. Модели принятия решений на основе лингвистической переменной. Рига. Зинатне. 1982.
- [**Душкин, 2011**] Душкин Р.В. Методы получения, представления и обработки знаний с НЕ-факторами. М. Вебов и Книгин. 2011.
- [**Куценко и др., 2008а**] Куценко Д.А., Синюк В.Г. Косвенный метод нечеткого вывода для продукционных систем со многими входами // Программные продукты и системы. 2008. № 1. С. 45–47.
- [**Куценко и др., 2008б**] Куценко Д.А., Синюк В.Г. Алгоритмы нахождения СР при кусочно-линейном представлении функций принадлежности // НСМВ-2008: сборник научных трудов второй Всероссийской научной конференции с международным участием. Ульяновск. УлГТУ. 2008. Т. 1. С. 87–92.
- [**Пегат, 2009**] Пегат А. Нечеткое моделирование и управление. М.: БИНОМ, Лаборатория знаний. 2009.
- [**Рутковский, 2010**] Рутковский Л. Методы и технологии искусственного интеллекта. М. Горячая линия – Телеком. 2010.
- [**Baczyński et al., 2008**] Baczyński M., Jayaram B. Fuzzy implications. Heidelberg. Springer. 2008.
- [**Dubois et al., 2012**] Dubois D., Prade H. Abe Mamdani: A Pioneer of Soft Artificial Intelligence // Combining Experimentation and Theory. Heidelberg. Springer. 2012. P. 49–60.
- [**Mamdani, 1974**] Mamdani E. H. Applications of fuzzy algorithm for control a simple dynamic plant // Proceedings of the IEEE. 1974. V. 121. No 12. P. 1585–1588.
- [**Zadeh, 1973**] Zadeh L. A. Outline of a new approach to the analysis of complex systems and decision processes // IEEE Transactions on Systems, Man and Cybernetics. 1973. V. SMC-3. No 1. P. 28–44.
- [**Zadeh, 1978**] Zadeh L. A. PRUF – A meaning representation language for natural language // International Journal of Man-Machine Studies. 1978. V. 10. P. 395–460.

УДК 004.89

**РАЗРАБОТКА МОДИФИКАЦИИ
АЛГОРИТМА RETE
ДЛЯ НЕЧЕТКИХ ЭКСПЕРТНЫХ СИСТЕМ**

И.С. Михайлов (*fr82@mail.ru*)

Национальный исследовательский университет Московский энергетический институт, Москва

Зо Мин Тайк (*zawgyi86@gmail.com*)

Национальный исследовательский университет Московский энергетический институт, Москва

В работе рассматриваются основные понятия нечетких экспертных систем. Предлагается разработанная модификация алгоритма Rete для нечеткой базы правил, позволяющая формулировать правила и заключения на ограниченном естественном языке и обеспечивающая ускорение процесса работы системы за счет однократного вычисления одинаковых условий в различных правилах.

Введение

В основе функционирования экспертных систем лежит модель знаний. Она содержит набор принципов, которые описывают состояние и поведение объекта исследования [ИИ, 1990]. Наиболее широко применяемой моделью знаний экспертных систем является продукционная модель в силу своей простоты обработки и понятности конечному пользователю [Вагин и др., 2001].

Однако в последнее время большое распространение приобретают нечеткие экспертные системы. Данный тип экспертных систем базируется на наборе правил, в которых используются лингвистические переменные и нечеткие отношения для описания состояния и поведения исследуемого объекта [Заде, 1976]. Правила, представленные в таком виде, наиболее приближены к естественному языку, поэтому нет необходимости использования отдельного специалиста инженера по знаниям для создания и редактирования правил. Они могут быть отредактированы самим экспертом практически без специальной подготовки [Гаврилова и др., 2001].

1. Постановка задачи

В рамках данной работы была поставлена задача разработать и реализовать модификацию алгоритма Rete для нечеткой продукционной базы правил. Также необходимо разработать нечеткую продукционную экспертную систему, механизм вывода которой будет функционировать на основе созданной модификации алгоритма Rete. Необходимо разработать формальную модель дерева решений модифицированного алгоритма Rete.

При работе решателя нечеткой продукционной экспертной системы все правила из базы правил применяются последовательно. При этом заново анализируются все условия, которые содержатся в данных правилах. Разрабатываемый алгоритм позволит проверять каждое условие только один раз. Таким образом, будет обеспечено ускорение работы решателя нечеткой ЭС и, следовательно, решение будет найдено быстрее.

2. Нечеткие экспертные системы

В основе правил работы нечетких экспертных систем лежит понятие лингвистической переменной. У каждой лингвистической переменной есть набор значений – это нечеткие переменные, которые образуют ее терм-множество. Лингвистическая переменная L характеризуется следующим набором свойств:

$$L = (X, T(X), U, G, M), \quad (1)$$

в котором: X – название переменной; $T(X)$ обозначает терм-множество переменной X , т.е. множество названий лингвистических значений переменной X , причем каждое из таких значений является нечеткой переменной x' со значениями из универсального множества U с базовой переменной u ; G – синтаксическое правило, порождающее названия x' значений переменной X ; M – семантическое правило, которое ставит в соответствие каждой нечеткой переменной x' ее смысл $M(x')$, т.е. нечеткое подмножество $M(x')$ универсального множества U .

Нечеткая переменная характеризуется тройкой $\langle x, U, X \rangle$, где x – название переменной, U – универсальное множество, X – нечеткое подмножество множества U , представляющее собой нечеткое

ограничение на значение переменной $u \in U$, обусловленное x [Заде, 1976].

Поведение исследуемой системы описывается на ограниченном естественном языке в терминах лингвистических переменных. Входные и выходные параметры системы рассматриваются как лингвистические переменные, а описание процесса задается набором правил. Формальная модель базы правил разработанной экспертной системы имеет вид:

$$\begin{aligned} L_1 : A_{11} \text{ и/или } A_{21} \text{ и/или } \dots \text{ и/или } A_{1m} &\rightarrow B_{11} \text{ и/или } \dots \text{ и/или } B_{1n}, \\ L_2 : A_{21} \text{ и/или } A_{22} \text{ и/или } \dots \text{ и/или } A_{2m} &\rightarrow B_{21} \text{ и/или } \dots \text{ и/или } B_{2n}, \\ &\dots \dots \dots \\ L_k : A_{k1} \text{ и/или } A_{k2} \text{ и/или } \dots \text{ и/или } A_{km} &\rightarrow B_{k1} \text{ и/или } \dots \text{ и/или } B_{kn}, \end{aligned} \quad (2)$$

где A_{ij} , $i = 1, 2, \dots, k, j = 1, 2, \dots, m$ – нечеткие высказывания, определенные на значениях входных лингвистических переменных, а где B_{ij} , $i = 1, 2, \dots, k, j = 1, 2, \dots, n$ – нечеткие высказывания, определенные на значениях выходных лингвистических переменных. Эта совокупность правил носит название нечеткой базы знаний.

На рисунке 1 представлена схема работы нечеткой экспертной системы.

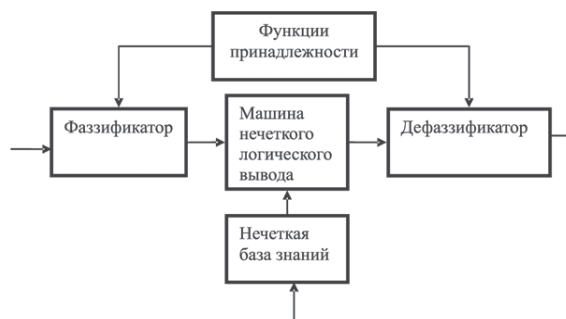


Рис. 1. Схема работы нечеткой экспертной системы

В общем случае нечеткий вывод решения происходит за три (или четыре) этапа:

1. Этап фаззификации. Преобразование с помощью функций принадлежности μ точных входных данных в нечеткие значения лингвистических переменных.

2. Этап непосредственного нечеткого вывода. На основании набора правил нечеткой базы знаний вычисляется значение истинности для условий каждого правила по правилам вычисления T-норм, T-конорм и отрицаний.

3. Этап композиции. Формируются значения выходных лингвистических переменных для каждого сработавшего правила.

4. Этап дефаззификации. Преобразование нечетких значений выходных лингвистических переменных в точные значения.

Для вычисления T-норм и T-конорм были выбраны следующие правила: $T_M(x, y) = \min\{x, y\}$ и $\perp_M(x, y) = \max\{x, y\}$. Для вычисления отрицания утверждения был выбран вариант классического отрицания $\neg(x) = 1 - x$.

3. Формальная модель дерева решений модифицированного Rete алгоритма

Алгоритм Rete является эффективным алгоритмом сопоставления с образцом для продукционных экспертных систем. Rete стал основой многих популярных экспертных систем, включая CLIPS, Jess, Drools, BizTalk Rules Engine и Soar. Классический алгоритм работы экспертной системы заключается в проверке применимости каждого правила вывода к каждому факту базы знаний, при необходимости выполнение его и переходе к следующему правилу с возвратом в начало базы знаний в случае исчерпания всех правил. Даже для небольшого набора правил и фактов такой метод работает неприемлемо медленно. Алгоритм Rete обеспечивает более высокую эффективность. При использовании Rete экспертная система строит специальный граф или префиксное дерево, узлам которого соответствуют части условий правил. Путь от корня до листа образует полное условие некоторой продукции. В процессе работы каждый узел хранит список фактов, соответствующих условию. При добавлении или модификации факта он прогоняется по сети, при этом отмечаются узлы, условиям которых данный факт соответствует. При выполнении полного условия правила, когда система достигает листа графа, правило выполняется.

В большинстве случаев скорость работы систем с использованием алгоритма Rete возрастает на порядки. Однако данный алгоритм реализован для экспертных систем с четкими продукцион-

ными правилами. В данной работе рассматривается модификация алгоритма Rete для работы с нечеткой базой правил.

Граф модификации Rete-алгоритма формируется аналогично четкому случаю, однако в каждом случае проверяется не точное значение условия правила, а значения лингвистических переменных в данном правиле. Формальная модель дерева решений будет иметь вид (3):

$$M = (X, R, P, Y), \quad (3)$$

где

1. X – множество вершин-условий графа. Каждая вершина представляет собой условия из правил нечеткой базы знаний.

$X = \{ \langle l_i, k_i, Tx_i, z_i \rangle \}$, где

- l_i – название лингвистической переменной;
- k_i – отношение сравнения, $k_i \in K$, где $K = \{=, \neq\}$;
- Tx_i – значение лингвистической переменной, то есть нечеткая переменная из ее терм-множества;
- z_i – значение степени уверенности μ , в том, что рассматриваемое условие l_i, k_i, Tx_i истинно. Данное значение вычисляется и означается при проверке рассматриваемого условия во время работы нечеткой экспертной системы;
- $i = 1 \dots n$, где n – количество условий в базе знаний.

2. Y – множество вершин-следствий графа. Каждая вершина представляет собой присваивание определенной лингвистической переменной ее значения. Каждая вершина представляет собой условия из правил нечеткой базы знаний:

$Y = \{ \langle l_i = Ty_i, z_i \rangle \}$, где

- l_i – название лингвистической переменной;
- Tx_i – значение лингвистической переменной, то есть нечеткая переменная из ее терм-множества;
- z_i – значение степени уверенности μ , в том, что рассматриваемое следствие $l_i = Tx_i$ истинно. Данное значение вычисляется и означается при определении степени уверенности в истинности правила, в которое входит данная лингвистическая переменная;
- $i = 1 \dots n$, где n – количество условий в базе знаний.

3. R – множество отношений между вершинами из объединения множеств множества X и Y . $R \subset ((X \cup Y) \times (X \cup Y))$. Или $R: (X \cup Y) \rightarrow (X \cup Y)$. R может принимать следующие значения: $R = \{ T, \perp, \top \}$.

В рассматриваемых далее T, \perp, \neg приняты следующие обозначения: x_i и x_j – условия из $(X \cup Y)$, $x_i \in (XUY)$, $x_j \in (XUY)$, z_i – значение степени уверенности в истинности условия $x_i = \langle l_i, k_i, Tx_i, z_i \rangle$, z_j – значение степени уверенности в истинности условия $x_j = \langle l_j, k_j, Tx_j, z_j \rangle$, z_k – значение степени уверенности в истинности результата вычисления нормы, конормы или отрицания.

- Отношение T-нормы определяется следующим образом: $z_k = T(x_i, x_j)$. Значения вычисляются по правилу:

$$T(x_i, x_j) = \min\{z_i, z_j\}.$$

- Отношение T-конормы определяется следующим образом: $z_k = \perp(x_i, x_j)$. Значения вычисляются по правилу:

$$\perp(x_i, x_j) = \max\{z_i, z_j\}.$$

- Отношение отрицания определяется следующим образом: $z_k = \neg(x_i)$. Значение вычисляется по правилу:

$$\neg(x_i) = 1 - z_i.$$

4. P – множество отношений, для описания правил нечеткой экспертной системы. $P \subset ((X \cup Y) \times Y)$. Или $P: (X \cup Y) \rightarrow Y$. В рассматриваемой модели P может принимать одно значение: $R = \{\rightarrow\}$, где \rightarrow связывает значения в левой части и вершины-следствия в правой части и означает, что при истинности значений в левой части (то есть при степени уверенности в них > 0.5) выполняются утверждения, которые расположены в правой части.

Далее рассмотрим процесс преобразования нечеткой базы правил в дерево решений, представленное в формате разработанной формальной модели.

4. Модификация алгоритма Rete для формирования дерева решений нечеткой продукционной базы правил

Данный алгоритм обрабатывает правила нечеткой базы правил и преобразует их в формат формальной модели дерева решений модифицированного алгоритма Rete. Рассмотрим схему данного алгоритма.

Входные данные алгоритма: правила из базы правил.

1. Начало алгоритма.
2. Устанавливаем $i = 1$.
3. Извлекаем из базы правил следующее непомеченное условие и формируем вершину x_i .

4. Просматриваем базу правил и помечаем все позиции в правилах, на которых встречается данное условие.

5. Формируем вершины для описания отношений из $R \{ T, \perp, \neg \}$, с которыми связано x_i . Устанавливаем отношения между данными вершинами.

6. Если в базе правил не закончились непомеченные условия, то увеличиваем i на 1 и переходим к шагу 3.

7. Устанавливаем $j=1$.

8. Извлекаем из базы правил следующее непомеченное следствие и формируем вершину y_i .

9. Просматриваем базу правил и помечаем все позиции в правилах, на которых встречается данное следствие.

10. Формируем вершину для описания отношения из $P \{ \rightarrow \}$, с которым связано y_j . Устанавливаем отношения между данными вершинами.

11. Если вершины И, ИЛИ, НЕ, с которыми связаны отношения из P , не сформированы, то формируем их.

12. Если вершины И, ИЛИ, НЕ, с которыми связаны отношения из P , уже сформированы, то устанавливаем связи с ними.

13. Если в базе правил не закончились непомеченные следствия, то увеличиваем j на 1 и переходим к шагу 8.

14. Конец алгоритма.

Выходные данные: дерево решений, построенное на основе правил.

В дальнейшем работа экспертной системы заключается в фаззификации входных данных и получении значений лингвистических переменных. Затем выполняется подстановка полученных значений в сформированное дерево решений, которое формирует результат работы экспертной системы. Далее может быть выполнена дефаззификация результата для получения числовых характеристик.

5. Сравнительный анализ предложенного алгоритма и существующих аналогов

Суть классического алгоритма Rete в том, что выделяются общие условия в различных правилах продукционной экспертной системы. Затем на их основе выполняется построение дерева реше-

ний. То есть каждая вершина алгоритма может быть либо истинна, либо ложна.

Модификация алгоритма Rete отличается от классического алгоритма тем, что он применяется для нечетких переменных. В связи с этим на каждом этапе работы алгоритма выполняется построение нечетких оценок истинности вершин дерева решений с помощью операторов Т-норм, Т-конорм и отрицания.

Это позволяет формулировать условия и следствия в базе правил на ограниченном естественном языке. Заключение формируются алгоритмом также на ограниченном естественном языке.

Вторым преимуществом является тот факт, что одинаковые условия объединяются при построении дерева решений на основе правил базы знаний. Это дает ускорение обработки дерева решений по сравнению с последовательным просмотром правил экспертной системы.

Оценка ускорения работы системы зависит от количества повторяющихся условий в различных правилах. При отсутствии повторяющихся условий модифицированный алгоритм Rete функционирует с такой же скоростью, как и классический алгоритм последовательной обработки правил. Однако в практических задачах однотипные условия встречаются достаточно часто. Задача, в которой будет получено максимальное ускорение работы системы с помощью применения модифицированного алгоритма Rete, содержит набор правил, в каждом из которых проверяется одинаковый полный набор из n условий, принимающих различные значения. При бинарном характере результатов проверки данных условий база будет содержать 2^n правил, и при классической схеме их просмотра понадобится проверить $n \cdot 2^n$ условий. При использовании модифицированного алгоритма Rete будет необходимо проверить лишь $2 \cdot n$ условий. В результате в предельном случае модифицированный алгоритм Rete позволит обрабатывать в 2^{n-1} раз меньше условий. Поэтому в качестве приблизительной оценки уменьшения количества проверяемых условий при формировании результата можно принять среднюю величину между двумя предельными случаями применения разработанного алгоритма, равную 2^{n-2} .

Модифицированный алгоритм Rete был реализован в рамках разработанной нечеткой экспертной системы диагностики заболеваний печени. Результаты тестирования на практике показали уско-

рение работы экспертной системы, функционирование которой построено на основе использования предложенного алгоритма, относительно скорости работы существующей экспертной системы.

Заключение

В работе получены следующие результаты:

1. Выполнено исследование основных понятий нечеткой логики и процесса создания и функционирования нечетких экспертных систем продукционного типа.

2. Изучены процессы и технологии разработки баз данных. Полученные навыки применены на практике при создании базы данных для тестовых примеров.

3. Разработана формальная модель дерева решений модифицированного алгоритма Rete для нечеткой продукционной базы знаний.

4. Предложена модификация алгоритма Rete для формирования дерева решений нечеткой продукционной базы правил.

5. Разработана экспертная система, которая позволяет формировать решение для подключаемой базы правил. Экспертная система преобразует нечеткую продукционную базу знаний в дерево решений Rete и в дальнейшем формирует заключение, на основе обработки данной структуры.

Разработанную нечеткую экспертную систему планируется применить в Московской медицинской академии им. Сеченова для диагностики у пациентов заболеваний печени.

СПИСОК ЛИТЕРАТУРЫ

[ИИ, 1990] Искусственный интеллект. В 3-х кн. Кн. 2. Модели и методы: справочник. – М.: Радио и связь, 1990. – 304 с.

[Вагин и др., 2001] Вагин В.Н., Еремеев А.П. Некоторые базовые принципы построения интеллектуальных систем поддержки принятия решений реального времени // Изв. РАН. ТиСУ, 2001, № 6, с. 114–123.

[Заде, 1976] Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 166 с.

[Гаврилова и др., 2001] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – Издательский дом «Питер», 2001. – 384 с.

УДК 004.891

ЭКСПЕРТНАЯ СИСТЕМА НА ОСНОВЕ НЕЧЕТКОЙ ЛОГИКИ И НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

Н.А. Полковникова (*npolkovnikova@tti.sfedu.ru*)

В.М. Курейчик (*kur@tgn.sfedu.ru*)

Южный федеральный университет,
Инженерно-технологическая академия (ИТА ЮФУ), г. Таганрог

В работе показана разработка модели нечеткой экспертной системы для идентификации неисправностей сложных технических объектов. Применение нейросетевых технологий позволяет обнаруживать нелинейные зависимости входных и выходных данных, повысить качество процесса диагностики объекта, что в конечном итоге позволит снизить количество аварийных ситуаций в условиях эксплуатации.

Введение

Проблема идентификации граничных состояний возникает при разработке экспертных систем (ЭС) для сложных технических объектов при распознавании критических ситуаций, связанных с неадекватной динамикой функционирования объекта контроля.

Основной трудностью при применении нечеткой логики в ЭС является необходимость явно сформулировать правила проблемной области в форме продукций. В этом случае эффективным решением является применение нейронных сетей, преимущество которых заключается в возможности автоматического переноса знаний лица, принимающего решение (ЛПР) в базу знаний (БЗ) ЭС. Внедрение нейросетевых технологий в ЭС или в существующие системы мониторинга при решении задач диагностики позволит не только фиксировать показания датчиков и сравнивать их с эталонными значениями, но и производить комплексный анализ получаемых параметров объекта, прогнозируя возможность наступления сбоев в работе как отдельных элементов, так и системы в целом [Покусаев, Касимов, 2012].

1. Структура ЭС

На рис. 1 представлена структурная схема гибридной ЭС, которая позволяет использовать преимущества нечеткой логики и нейронных сетей. БЗ в ЭС представлена в форме нечетких про-

дукций и лингвистических переменных. Основу лингвистической переменной составляет терм с функцией принадлежности.

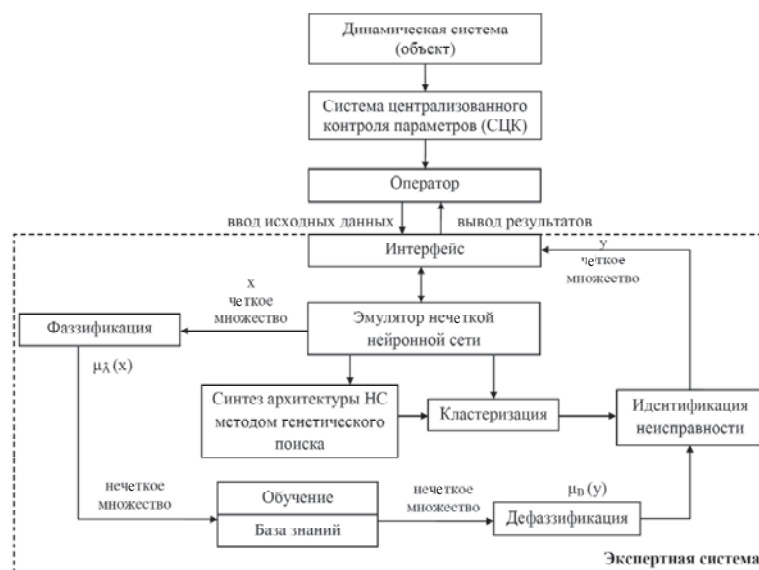


Рис. 1. Структурная схема взаимодействия оператора с экспертной системой

Использование аппарата нечеткой логики при разработке БЗ и механизмов вывода ЭС позволяет формализовать процедуру оценки технического состояния на базе фрагментарной, ненадежной и, возможно, неточной информации и обоснованно принимать решения по идентификации неисправностей.

Гибридная ЭС извлекает знания из статистических выборок, интерпретируя их как обучающие выборки для обучения нечеткой нейронной сети [Ярушкина, 2004]. БЗ гибридной ЭС содержит следующие компоненты: функции принадлежности, нечеткие продукции, обученные нечеткие нейронные сети. Ретроспективная информация, хранящаяся в базах данных, служит основой при нахождении шаблонов, отражающих динамику поведения входных параметров для идентификации граничных состояний с помощью нейронных сетей.

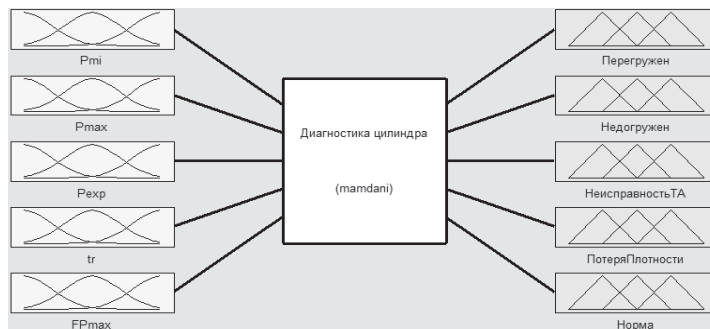


Рис. 2. Функциональная схема идентификации неисправностей цилиндра двигателя в окне редактора MATLAB

На рис. 2 представлена функциональная схема диагностики цилиндра судового двигателя в окне редактора системы нечеткого вывода. Входами являются значения параметров цилиндра из табл. 1, полученные в результате индицирования. Выходами являются вероятные неисправности цилиндра, характерные для данных входных параметров. Цилиндр двигателя может иметь следующие состояния: перегружен, недогружен, неисправность топливной аппаратуры (ТА), потеря плотности, норма [Полковникова, Курейчик, 2014].

2. Фаззификация

На первом этапе формализации нечеткого логического вывода является фаззификация входных параметров (рис. 1). Процесс задания нечеткого множества на основе известного количественного значения признака называется фаззификацией или приведением к нечеткости. Фаззификация – это также процесс нахождения значений функций принадлежности нечетких множеств на основе четких исходных данных.

Фаззификация позволяет представить объективно присутствующую неточность результатов физических измерений. На этапе фаззификации значения входных параметров в виде N-мерного входного вектора $x = [x_1, x_2, \dots, x_n]^T$ приводятся в нечеткое множество \tilde{A} в соответствии с их лингвистической оценкой и последующим выбором закона изменения функции принадлежности $\mu_{\tilde{A}}(x)$.

Таблица 1

Значения параметров в каждом диапазоне

Наименование параметра	Обозн.	Значение параметров		
		низкое	нормальное	высокое
Среднее индикаторное давление, бар	P_{mi}	[15...15,6]	[15,6...16,4]	[16,4...17]
Максимальное давление в цилиндре, бар	P_{max}	[130...141,5]	[141,5...148,5]	[148,5...160]
Давление на линии расширения при угле 36° за ВМТ, бар	P_{exp}	[55...63]	[63...67]	[67...75]
Температура выпускных газов, $^\circ C$	t_r	[270...285]	[285...315]	[315...330]
Максимальное давление в топливопроводе высокого давления, бар	P_{Pmax}	[500...520]	[520...580]	[580...600]

При разработке БЗ ЭС для главного судового двигателя использованы результаты индицирования (табл.1), которые преобразованы в три диапазона входных значений параметров соответственно для трех лингвистических переменных (термов): low (низкое), normal (нормальное), high (высокое). Для каждого входного параметра был задан трапециевидный закон изменения функции принадлежности: $y = \text{trapmf}(x, [a, b, c, d])$, где аргумент a – минимальное допустимое значение параметра с нулевой вероятностью нормы, отрезок между аргументами $[b, c]$ показывает принадлежность параметра к норме с вероятностью 1. И, соответственно, аргумент d – максимальное допустимое значение параметра с нулевой вероятностью нормы. Функции принадлежности являются нормальными (нормированными), принимающими значения в интервале от 0 до 1. Выбор трапециевидной формы функции принадлежности обусловлен тем, что принадлежность параметра к норме с вероятностью 1 определяется не одним-единственным значением переменной, а некоторым диапазоном из допустимых значений.

3. Дефаззификация

Вторым этапом формализации нечёткого логического вывода является дефаззификация (рис. 1). Дефаззификатор трансформирует нечеткое множество в полностью детерминированное точечное решение y . Нечеткое множество представляет зависимость $\mu(y) = \mu_{A \rightarrow B}(y)$ как функцию от выходной переменной y . Результат $\mu(y)$ состоит из суммы нечетких функций для имплика-

ции всех M правил, образующих систему нечеткого вывода. В модели Мамдани-Заде каждое из этих M правил определяется уровнем активации условия, $\mu(y_i) = \prod_{j=1}^N \mu_A(x_j)$, тогда как y_i – это значение y , при котором величина $\mu_B(y)$ становится максимальной (либо принимает среднее из максимальных значений).

4. Формализация БЗ

Следующим этапом является формализация БЗ. Принципиально БЗ можно построить, используя либо нечеткий логический вывод по Мамдани-Заде, либо вывод Такаги-Сугено-Канга (ТСК) с помощью нейронной сети. Отличие базы знаний ТСК от Мамдани-Заде заключается в том, что заключения правил задаются не нечеткими термами, а функцией от входов. Однако БЗ Мамдани-Заде можно построить, прибегнув к знаниям ЛПР, а БЗ ТСК целесообразно использовать в случае отсутствия у оператора (ЛПР) необходимых знаний об объекте [Полковникова, Курейчик, 2013].

Разработка структуры нейронной сети по нечеткой адаптивной модели вывода ТСК для идентификации граничных состояний была произведена с помощью ANFIS-сети (ANFIS – Adaptive Network Based Fuzzy Inference System – адаптивная сеть нечеткого вывода). ANFIS-сеть представляет собой систему нечеткого вывода ТСК нулевого или первого порядка, в которой каждое из правил нечетких продукций имеет постоянный вес, равный 1. При

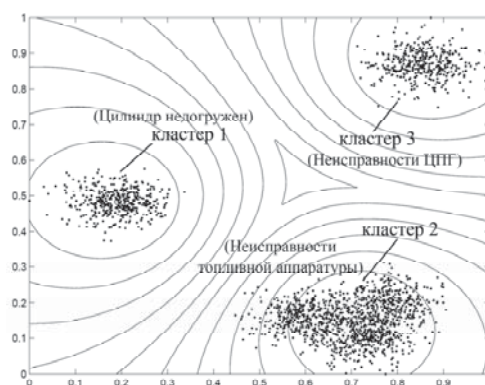


Рис. 3. Результат кластеризации для определения технического состояния цилиндра

построении нейронной сети сначала синтезируются нечеткие правила из экспериментальных данных с использованием кластеризации. После кластеризации настраиваются параметры ANFIS-сети. Кластеризация в отличие от классификации не имеет заранее заданных групп. С помощью кластеризации нейронная сеть самостоятельно выделяет различные однородные неисправности в группы (кластеры) на основе схожести признаков (рис. 3).

Кластеризация необходима для сокращения объема информации и области поиска, что позволяет сократить время для идентификации неисправности, а также учета динамики изменения результатов индицирования двигателя во времени.

ANFIS-сеть реализует систему нечеткого вывода Сугено в виде пятислойной нейронной сети, входы которой в отдельный слой не выделяются. На рис. 4 представлена разработанная нейронная ANFIS-сеть для идентификации неисправности топливной аппаратуры главного судового двигателя. Каждый узел первого слоя представляет один терм с трапецевидной функцией принадлежности. Выходом узла является степень принадлежности значения

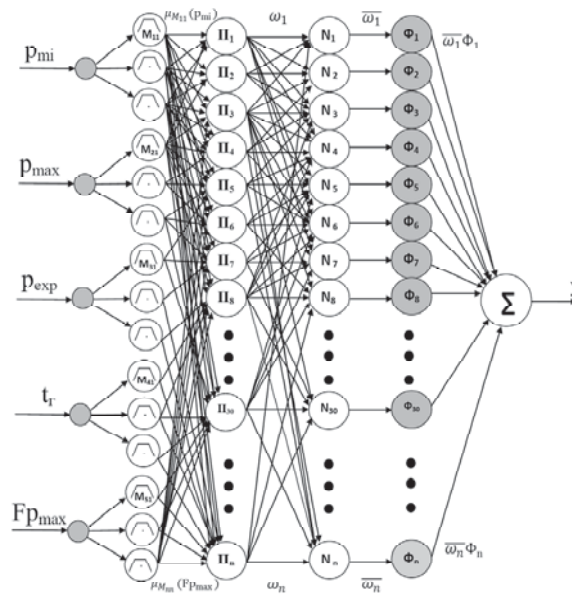


Рис. 4. Структура адаптивной нейронечеткой ANFIS-сети

входной переменной соответствующему нечеткому терму. Каждый узел второго слоя соответствует одному нечеткому правилу. Узел второго слоя соединен с теми узлами первого слоя, которые формируют antecedentes соответствующего правила. Выходом узла является степень выполнения правила, которая рассчитывается как произведение входных сигналов. Задача нейронной сети состоит в таком отображении входных данных, чтобы ожидаемое значение, соответствующее входному вектору $P_{mi}, P_{max}, P_{exp}, t_2, FP_{max}$, формировалось выходной функцией сети y . В качестве исходных данных используются результаты индизирования двигателя, соответствующие определенному периоду эксплуатации при неисправности топливной аппаратуры.

Обучение нейронной сети сводится к такой настройке параметров функций принадлежности, при которой обеспечивается минимальное значение ошибки обучения. Для нейронной сети был выбран алгоритм обратного распространения ошибки, поскольку он считается одним из наиболее эффективных алгоритмов обучения многослойных сетей. Его основу составляет целевая функция, формулируемая в виде квадратичной суммы разностей между фактическими и ожидаемыми значениями выходных сигналов. В качестве исходных данных используются результаты индизирования двигателя, соответствующие определенному периоду эксплуатации при неисправности топливной аппаратуры. Обучение нейронной сети сводится к такой настройке параметров функций принадлежности, при которой обеспечивается минимальное значение ошибки обучения. Для нейронной сети был выбран алгоритм обратного распространения ошибки, поскольку он считается одним из наиболее эффективных алгоритмов обучения многослойных сетей. Его основу составляет целевая функция, формулируемая в виде квадратичной суммы разностей между фактическими и ожидаемыми значениями выходных сигналов. При нескольких обучающих выборках j ($j = 1, 2, \dots, P$), целевая функция определяется в виде суммы по всем выборкам:

$$E(w) = \frac{1}{2} \sum_{j=1}^P \sum_{r=1}^R (y_j^r - d_j^r)^2,$$

где y_j^r – фактическая реакция нейронной сети на входной вектор

$$X^r, X^r = (P_{mi}^r, P_{max}^r, P_{exp}^r, t_2^r, FP_{max}^r)^T;$$

d^r – желаемое значение выхода нейронной сети;

R – размерность обучающей выборки (количество обучающих пар y, d).

Ошибка обучения нейронной сети за 50 итераций (эпох) составила $2,3 \times 10^{-6}$. Проверка построенной модели гибридной ANFIS-сети выполнена по результатам нечеткого вывода для этих значений входных переменных, полученных с помощью разработанной нечеткой модели по Мамдани-Заде. По результатам нечеткого логического вывода получены поверхности «входы-выход», отображающие динамику развития неисправности топливной аппаратуры (рис. 5).

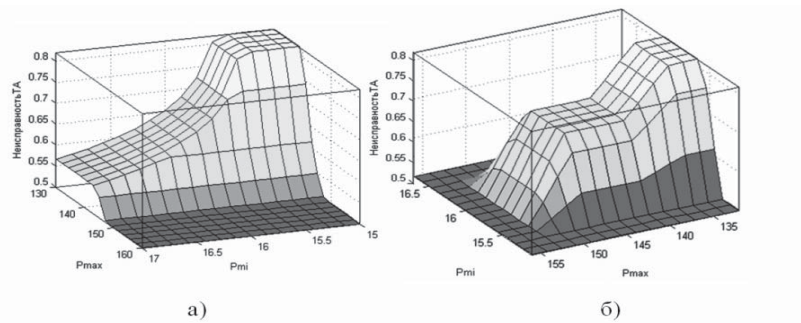


Рис. 5. Определение вероятности неисправности топливной аппаратуры: а) – по Мамдани-Заде, б) – по ТСК с использованием ANFIS-сети

Поверхности отображают вероятность неисправности топливной аппаратуры при изменении входных параметров. Вероятность неисправности топливной аппаратуры достигает максимального значения $0,75 \dots 1$ при параметрах: $p_{mi} = 15$ бар, $p_{max} = 131$ бар, $p_{exp} = 71$ бар, $t_r = 325$ °C и $Fp_{max} = 505$ бар. Проверка показывает согласованность нейронной сети и подтверждает ее адекватность в рамках рассматриваемой модели.

Заключение

Использование аппарата нечеткой логики при разработке ЭС позволяет формализовать процедуру оценки технического состояния на базе фрагментарной, ненадежной и, возможно, неточной ин-

формации и с высокой степенью вероятности идентифицировать граничные состояния сложных технических объектов. Применение нейросетевых технологий в ЭС позволяет эффективно аппроксимировать как линейные, так и сложные нелинейные зависимости, повысив качество процесса диагностики объекта, что в конечном итоге позволит существенно снизить количество аварийных ситуаций.

В работе показана реализация модели ЭС на основе нечеткой логики и нейросетевых технологий для идентификации неисправностей главного судового дизельного двигателя на основе параметров рабочего процесса и топливоподачи. Такая модель позволяет на ранней стадии адекватно отразить растущую вероятность неисправности при изменении входных параметров в режиме реального времени.

СПИСОК ЛИТЕРАТУРЫ

[Полковникова, Курейчик, 2014] Полковникова Н.А., Курейчик В.М. Разработка модели экспертной системы на основе нечеткой логики // Известия ЮФУ. Технические науки. – 2014. – №1 (150) – С. 83–92.

[Покусаев, Касимов, 2012] Покусаев М.Н., Касимов Н.Н. Система диагностики судовых энергетических установок с применением нейросетевых моделей // Вестник АГТУ. Сер.: Управление, вычислительная техника и информатика, №3, 2012. – С. 88–93.

[Полковникова, Курейчик, 2013] Полковникова Н.А., Курейчик В.М. Об интеллектуальном анализе баз данных для экспертной системы // Информатика, вычислительная техника и инженерное образование. – 2013. – №2 (13) – С. 39–50.

[Осипов и др., 2012] Осипов Г.С., Толмачев И., Хачумов В. Принципы построения системы мониторинга, моделирования и прогнозирования чрезвычайных ситуаций // Фундаментальные проблемы безопасности. 2012. № 3. С. 92–97.

[Ярушкина, 2004] Ярушкина Н.Г. Основы теории нечетких и гибридных систем: Учеб. пособие. – М.: Финансы и статистика, 2004. – 320 с.

УДК 004.89:53.08:656.2

СИСТЕМЫ ПРИОБРЕТЕНИЯ ЗНАНИЙ ТРЕТЬЕГО ПОКОЛЕНИЯ НА ОСНОВЕ КОГНИТИВНЫХ ИЗМЕРЕНИЙ¹

М.Н. Святкина (*maria.svyatkina@gmail.com*)

МГТУ им. Н.Э. Баумана, Москва

В.Б. Тарасов (*tarasov@rk9.bmstu.ru*)

МГТУ им. Н.Э. Баумана, Москва

Проанализированы модели неклассических измерений для систем приобретения знаний третьего поколения. Предложена новая концепция когнитивных измерений как процесса иерархической грануляции измерительной информации на основе принципа единства измерений и прагматических оценок. Рассмотрена двухуровневая архитектура когнитивных измерений. Разработан логико-алгебраический подход к интерпретации и анализу данных на верхнем уровне когнитивных измерений.

Введение

Первое поколение систем приобретения знаний было связано с извлечением знаний из уникального эксперта (экстенсивная схема приобретения знаний) и наполнением «оболочек» экспертных систем. Второе поколение, предполагающее разработку системы онтологий и проведение онтологического инжиниринга, обеспечивает построение концептуальных моделей, разделяемых сообществом специалистов. Прежде всего, оно ориентировано на достижение взаимопонимания и осуществление совместной работы агентов в многоагентных системах. Наконец, в интегрированных системах приобретения знаний 3-го поколения наряду с онтологиями широко используются средства получения и интеграции информации от сенсоров и ее последующей обработки с помощью специального процессора интеллектуального анализа данных и обнаружения знаний.

В работе предлагается новая концепция когнитивных измерений на основе гранул, построены гранулярные прагматики для отдельных сенсоров и их диад.

¹ Работа выполнена при финансовой поддержке РФФИ (проект №14-07-31317).

1. Концепции неклассических измерений

В конце XX – начале XXI века появились и сегодня активно развиваются новые направления неклассических измерений: автономные измерения, распределенные измерения на основе сенсорных сетей, интеллектуальные [Finkelstein, 1994], [Романов и др., 1994], мягкие [Аверкин и др., 1997], когнитивные [Прокопчина, 2010] измерения и пр. При этом ощущается заметное влияние новых концепций вычислений на методологию измерений. Так, в 1994 г. Л.Заде опубликовал статью о «мягких вычислениях» (Soft Computing) [Zadeh, 1994], а три года спустя А.Н. Аверкин и С.В.Прокопчина [Аверкин и др., 1997] по аналогии ввели понятие «мягких измерений» (Soft Measurements). В свою очередь, концепция грануляции информации Л.Заде [Zadeh, 1997] и «гранулярные вычисления» (Granular Computing) Т.Лина [Lin, 1998] дали импульс «гранулярным измерениям» [Тарасов, 2013].

Термин «когнитивные измерения» был предложен С.В. Прокопчиной в статье [Прокопчина, 2010], где речь идет об измерениях, в результате которых с помощью байесовских интеллектуальных технологий извлекаются и используются метрологически аттестованные знания.

В настоящей работе развивается концепция когнитивных измерений как процесса иерархической грануляции информации на основе принципа единства измерений и прагматических оценок. Ей соответствует двухуровневая архитектура измерений (рис. 1), где на нижнем уровне с помощью набора датчиков, осуществляющих количественные измерения, добывается мелкозернистая информа-

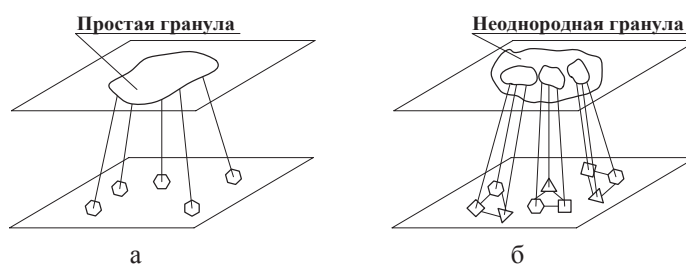


Рис. 1. Иллюстрация формирования простых и составных гранул:
а) грануляция данных от однотипных датчиков;
б) грануляция данных от разнотипных датчиков

ция, а на верхнем уровне результаты этих измерений отображаются в прагматическую шкалу лингвистических значений типа «норма», «предотказ 1-го рода», «предотказ 2-го рода», «полный отказ» и т.п. (крупнозернистая информация).

Будем рассматривать гранулярные измерения как концептуальное ядро когнитивных и мягких измерений и как материал для интеллектуального анализа сенсорных данных с выявлением закономерностей в этих данных.

2. Гранулы, процесс грануляции и гранулярные измерения

В современной теории измерений неявно полагается, что и полученное значение измеряемой величины, и его неопределенность представляют собой гранулы. Реализация гранулярных измерений предполагает построение теории грануляции и развитие теории измерений. Грануляция включает создание, интерпретацию и представление гранул, тогда как развитие теории измерений в данном контексте связано с единым пониманием измерений как иерархических гранулярных структур.

Вначале определим термины: гранула, грануляция, гранулярная структура. Согласно Л.Заде [Zadeh, 1997], гранулой называется группа объектов, объединяемых отношениями неразличимости, сходства, близости или функциональности, т.е. отношениями, обладающими, по крайней мере, свойствами рефлексивности и симметричности. Грануляция – способность представлять и оперировать данными, информацией, знаниями на различных уровнях детализации – представляется одним из ключевых свойств когнитивных измерений. Здесь термин «Грануляция» охватывает процессы композиции (формирование более крупных гранул) и декомпозиции (формирование более мелких гранул). Выбирая различные степени грануляции измерительной информации, определяем различные уровни детализации знаний. Гранулярная структура – это семейство гранул, связанных между собой отношением вложенности.

В основе гранулярных измерений лежит формирование простых и составных, однородных и неоднородных, одномерных и многомерных гранул на основе полиморфных соотношений между первичными измерениями (нижний уровень) и их логико-лингвистическими оценками (верхний уровень).

К числу фундаментальных проблем теории грануляции измерений относятся: определение общих принципов и критериев представления результатов измерений в форме гранул; анализ способов интерпретации и классификации гранул; построение формальных моделей гранул и гранулярных структур; разработка методов и процедур грануляции информации в процессе измерений; установление соотношений между гранулами и уровнями грануляции; определение количественных характеристик гранул и процесса грануляции.

При построении системы гранулярных измерений используем принцип грануляции Л.Заде: степень грануляции измерительной информации должна соответствовать допустимому уровню неточности при решении конкретной задачи. Критерии грануляции тесно связаны с прагматической интерпретацией гранул и должны давать ответ на вопрос, почему несколько измеренных значений включены в одну и ту же гранулу.

Информационно-измерительные гранулы различаются по своей природе, размеру, сложности, неоднородности и т.п. Характерные интерпретации гранул в измерениях – это часть целого, единица знания, область неопределенности. В теории измерений можно выделить интервальные, вероятностные и нечеткие, простые и составные, однородные и неоднородные гранулы. Классическими гранулярными понятиями являются плотность вероятности и доверительный интервал.

В случае нечетких измерений термин «гранулы» охватывает нечеткие числа и нечеткие интервалы, распределения возможности, уверенности и правдоподобия, нечеткие отображения и лингвистические метки, четкие и нечеткие множества уровня, и пр. Ниже изложим логико-алгебраический подход к интерпретации и анализу данных на верхнем уровне когнитивных измерений.

3. Когнитивные сенсоры: использование многозначных логик с гранулярной прагматикой в интерпретации измерений

В основе когнитивных измерений лежат когнитивные сенсоры, которые мы определяем с помощью многозначных логик (см., например, [Карпенко, 2010], [Финн, 2008]) и семиотики Ч. Пирса [Пирс, 2000]. С одной стороны, логические модели грануляции опи-

раются на многозначные и паранепротиворечивые логики. Родоначальником этих логик был русский ученый Н.А.Васильев, который построил двухуровневую логическую структуру (эмпирическая логика и металогика), а также ввел многомерные (воображаемые) логики [Васильев, 1989] (иллюстрацией трехмерных логических высказываний является треугольник Васильева – см. рис. 2). По аналогии мы разбиваем двухуровневую структуру когнитивных измерений на эмпирический уровень и уровень метаизмерений (прагматических оценок полученных количественных величин).

Общую стратегию формирования гранулярной логической семантики предложил Дж.Данн. Пусть V – множество значений истинности. По Данну, значениями истинности могут быть не только отдельные элементы $v \in V$, но и любые его подмножества, а также пустое множество \emptyset , т.е. производится изящный переход от V к 2^V [Dunn, 1976].

В простейшем случае для $|V|=2$ получаем логическую семантику Данна-Белнапа $2^V = V_4 = \{T, B, N, F\}$, где $B = \{T, F\}$, $N = \{\emptyset\}$. Здесь B (от англ. слова Both – и то, и другое, оба) понимается как «пресыщенная оценка» истинности (Glut), а N (от None – ни то, ни другое) – как разрыв истинности или «истиннозначный провал» (Gap). На рис. 3а и 3б приведены диаграммы Хассе для логической решетки Белнапа [Belnap, 1977] и двойственной ей решетки Скотта.

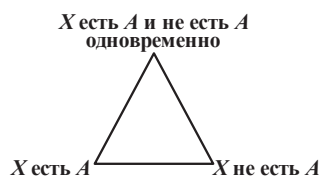


Рис. 2. Треугольник Васильева

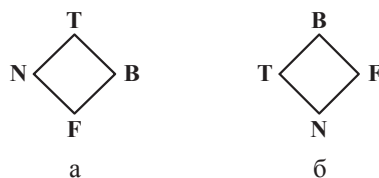


Рис. 3. Диаграммы Хассе для решеток Белнапа (а) и Скотта (б)

С другой стороны, ниже развивается прагматический подход к когнитивным измерениям, опирающийся на идеи Ч.Пирса [Пирс, 2000]. Им была предложена концепция семиотики как «логики в самом общем смысле», формальной доктрины знаковых систем. В соответствии с прагматической максимой Ч.Пирса, мы постигаем предмет нашей мысли, рассматривая те его свойства, которые несут большее практическое значение. Здесь имеется в виду прагматика

измерений, связанная с пониманием и полезностью результатов измерений в плане решения конкретных задач диагностики, прогнозирования, мониторинга и пр.

Будем связывать с каждым результатом измерения m из множества M одну (или несколько) оценку его истинности $v \in V$, где истинность трактуется в русле прагматической теории как практическая полезность измеренного значения для решения какой-либо задачи.

3.1. Васильевские и белнаповские сенсоры

Будем понимать под когнитивным сенсором любой датчик, способный не только измерять значения некоторого параметра рассматриваемого объекта, но и интерпретировать (понимать) полученную информацию. Для оценки и интерпретации численных показаний датчика можно использовать трехзначную логическую прагматику. Соответствующий сенсор будем называть васильевским в честь Н.А. Васильева. В нем полученная количественная информация гранулируется по трем прагматическим значениям («светофорный принцип»): 1) T – «прагматическая истина» («параметр в норме» – показания сенсора находятся в «зеленой зоне»); 2) F – «прагматическая ложь» («параметр не в норме», «отказ» – показания сенсора попали в «красную зону»); 3) B – «прагматическое противоречие» («пограничная ситуация» или «предотказ» – показания сенсора локализованы в «желтой зоне»). Если в дополнение к указанным трем значениям мы также рассмотрим: 4) N – «полная неопределенность» (ресурсы сенсора истощены и нет никаких показаний, или датчик «спит»), то переходим к белнаповскому сенсору. Работу такого сенсора можно описать логической матрицей

$$LM_{V_4} = \langle \{T, B, N, F\}, \{\neg, \wedge, \vee, \rightarrow\} \{F\} \rangle, \quad (1)$$

где T, B, N, F интерпретируются соответственно как прагматическая истина, прагматическое противоречие, полная неопределенность, прагматическая ложь, $\neg, \wedge, \vee, \rightarrow$ есть основные логические операции над значениями истинности (отрицание, конъюнкция, дизъюнкция, импликация), а F – антивыведенное значение, поскольку в задаче мониторинга объектов инфраструктуры нас интересует диагностирование неработоспособного состояния «отказ».

Следует отметить, что в реальной жизни границы между зеленой, желтой и красной зонами показаний сенсоров являются размы-

тыми и динамичными. Кроме того, при анализе мультисенсорной информации мы можем столкнуться с ситуациями, выражаемыми качественно в виде «данные от большинства сенсоров находятся в зеленой зоне» или «данные нескольких сенсоров сигнализируют о возможном предостережении». В таких случаях требуется переход от обычных к нечетким вазильевским и белнаповским сенсорам, снабженным соответствующей прагматикой. Для белнаповского сенсора такую прагматику можно выразить четверкой

$$V_f(m) = \{T_f(m), B_f(m), N_f(m), F_f(m)\}, \quad (2)$$

где m есть результат измерения, а T_f, B_f, N_f, F_f – нечеткие значения степени истины, противоречия, неопределенности и лжи соответственно, $T_f, B_f, N_f, F_f \in [0,1]$. Вариант нечеткой логики Белнапа описан в [Тугупев и др., 2010].

3.2. Гранулярные прагматики мультисенсорных данных

Для построения логик сенсорных сетей будем использовать произведение логик в виде произведения соответствующих решеток. Логическая прагматика сети из n белнаповских сенсоров будет описываться формулой 4^n , где n – целое число, $n > 1$. Простейшая нетривиальная сеть белнаповских сенсоров образует множество прагматических значений $4^2=16$, сеть из трех датчиков – $4^3=64$ и т.д.

С целью иллюстрации предлагаемого подхода возьмем задачу интерпретации показаний двух белнаповских сенсоров, моделируемую с помощью бирешеток [Ginsberg, 1988].

Бирешеткой называется четверка

$$BL = \langle U, \leq_1, \leq_2, \neg \rangle, \quad (3)$$

где $U = X \times X \neq \emptyset$, $|U| \geq 2$, \leq_1 и \leq_2 – два различных отношения порядка, заданных на множестве U , а $\neg: U \rightarrow U$ есть отрицание Гинзберга, которое удовлетворяет следующим условиям: $\forall x, y \in U$, 1) если $x \leq_1 y$, то $\neg x \geq_1 \neg y$; 2) если $x \leq_2 y$, то $\neg x \leq_2 \neg y$; 3) $\neg(\neg x) = x$.

Очевидно, что бирешетка (3) может строиться как алгебра с двумя различными операциями пересечения и объединения соответственно, т.е. задаваться шестеркой:

$$BL = \langle X \times X, \wedge, \vee, \otimes, \oplus, \neg \rangle, \quad (3^*)$$

где: 1) решетки $\langle X, \wedge, \vee \rangle$ и $\langle X, \otimes, \oplus \rangle$ – полные; 2) \neg есть отображение $\neg: X \rightarrow X$, такое что (а) $\neg^2 = 1$; (б) \neg есть гомоморфизм решетки $\langle X, \wedge, \vee \rangle$ в решетку $\langle X, \vee, \wedge \rangle$ и автоморфизм решетки $\langle X, \otimes, \oplus \rangle$.

Все конечные бирешетки можно представить графически с помощью двойных диаграмм Хассе, где по вертикали (снизу вверх) отображается один порядок, а по горизонтали (слева направо) – другой порядок. Здесь классическим примером служит минимальная бирешетка «4» с порядком истинности \leq , и информационным порядком \leq_i (рис. 4). На рис. 5 приведена диаграмма Хассе бирешетки «9» = 3^2 , для диады васильевских сенсоров.

Рассмотрим все прагматические значения для диад сенсоров. Здесь T_1T_2 – «согласованная прагматическая истина» (показания обоих сенсоров определяются значением «норма»); F_1F_2 – «согласованная прагматическая ложь» (оба сенсора сигнализируют о состоянии отказа); $T_1B_2 \sim B_1T_2$ – «частичное прагматическое противоречие как предотказ 1-го рода» (один сенсор показывает значение «норма», а другой – «предотказ»); $T_1N_2 \sim N_1T_2$ – «частичная прагматическая истина с неопределенностью» (один сенсор показывает значение «норма», а другой «спит»); $T_1F_2 \sim F_1T_2$ – «полное прагматическое противоречие» (один сенсор показывает значение «норма», а другой – «отказ»); B_1B_2 – «согласованное прагматическое противоречие» (оба сенсора передают состояние «предотказ»); N_1N_2 – «согласованная неопределенность» (оба сенсора исчерпали свои ресурсы или оба сенсора «спят»); $F_1B_2 \sim B_1F_2$ – «частичное прагматическое противоречие как предотказ 2-го рода» (один сенсор показывает значение «отказ», а другой – «предотказ»); $F_1N_2 \sim N_1F_2$ – «частичная прагматическая ложь с неопределенностью» (один сенсор показывает значение «отказ», а другой спит).

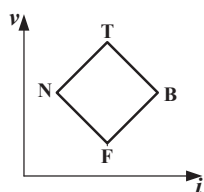


Рис. 4. Двойная диаграмма Хассе для бирешетки «4»

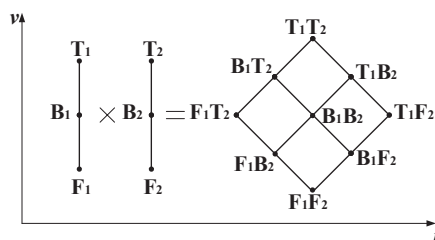


Рис. 5. Двойная диаграмма Хассе для бирешетки «9», иллюстрирующей логическую прагматику информации от двух васильевских сенсоров

Из рис. 6 видно, что, например, $T_1T_2 >_v T_1B_2 >_v T_1F_2 >_v B_1F_2 >_v F_1F_2$.

В логике Белнапа значения B и N рассматриваются независимо друг от друга, поэтому при ее консервативном расширении значения B_1N_2 и N_1B_2 запрещены. Таким образом, $|V|=14$, т.е. имеем четырнадцатизначную прагматику диалога сенсоров (рис. 6). Общая карта отказов и предотказов дана на рис. 7.

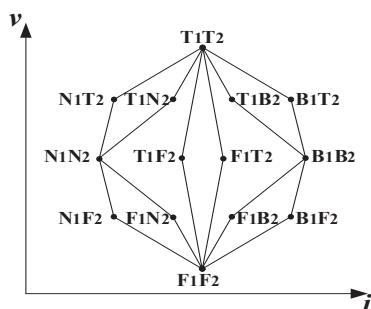


Рис. 6. Бирешеточное представление прагматики информации от двух белнаповских сенсоров

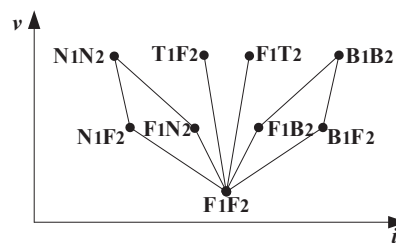


Рис. 7. Совмещенная карта отказов и предотказов для двух белнаповских сенсоров

Заключение

В работе предложена новая концепция когнитивных измерений как процесса иерархической грануляции информации в интересах построения интеллектуальных систем новых поколений. В ней использованы идеи логической прагматики Пирса, трехзначные и четырехзначные логики и их произведения, элементы теории бирешеток. Дальнейшие перспективы исследования связаны с соединением подходов гранулярных измерений и интеллектуального анализа данных (Data Mining), развитием методов комплексирования сенсорных данных и обнаружения знаний, в том числе с помощью гранулярных логик на базе трирешеток и квадранеток.

СПИСОК ЛИТЕРАТУРЫ

- [Аверкин и др., 1997] Аверкин А.Н., Прокопчина С.В. Мягкие вычисления и измерения // Интеллектуальные системы (МГУ). – 1997. – Т. 2, вып. 1–4.
- [Васильев, 1989] Васильев Н.А. Воображаемая логика. – М.: Наука, 1989.

- [Карпенко, 2010]** Карпенко А.С. Развитие многозначной логики. – М.: Изд-во ЛКИ, 2010.
- [Пирс, 2000]** Пирс Ч.С. Избранные произведения: Пер. с англ. – М.: Логос, 2000.
- [Прокопчина, 2010]** Прокопчина С.В. Когнитивные измерения на основе байесовских интеллектуальных технологий// Сборник докладов XIII Международной конференции по мягким вычислениям и измерениям (SCM'2010, Санкт-Петербург, СПбГЭТУ «ЛЭТИ», 23–25 июня 2010 г.). – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2010.
- [Романов и др., 1994]** Романов В.Н., Соболев В.С., Цветков Э.И. Интеллектуальные средства измерения/ Под ред. Э.И. Цветкова. – М.: Татьянин день, 1994.
- [Тарасов, 2013]** Тарасов В.Б. О гранулярных структурах измерений в интеллектуальных средах: васильевские и белнаповские сенсоры и модели их взаимодействия// Информационно-измерительные и управляющие системы. – 2013. – Т. 11, №2.
- [Финн, 2008]** Многозначные логики и их применения. В 2-х томах/ Под ред. В.К.Финна. – М.: Изд-во ЛКИ, 2008.
- [Belnap, 1977]** Belnap N. A Useful Four-Valued Logic// Modern Uses of Multiple-Valued Logic/ Ed. by J.M.Dunn and G.Epstein. – Dordrecht: D.Reidel, 1977.
- [Dunn, 1976]** Dunn J.M. Intuitive Semantics for First-Degree Entailment and «Coupled Trees»// Philosophical Studies. – 1976. – Vol.29.
- [Finkelstein, 1994]** Finkelstein L. Intelligent and Knowledge-Based Instrumentation. An Examination of Basic Concepts// Measurement. – 1994. – Vol.14.
- [Ginsberg, 1988]** Ginsberg M. Multi-Valued Logics: a Uniform Approach to Reasoning in AI// Computer Intelligence. – 1988. – Vol.4.
- [Lin, 1998]** Lin T.Y. Granular Computing on Binary Relations I: Data Mining and Neighborhood Systems// Rough Sets in Knowledge Discovery/ Ed. by A. Skowron and L. Polkowski. – Heidelberg: Physica-Verlag, 1998.
- [Turunen и др., 2010]** Turunen E., Ozturk M., Tsoukias A. Paraconsistent Semantics for Pavelka Style Fuzzy Sentential Logic. Fuzzy Sets and Systems. – 2010. – Vol.161, №14.
- [Zadeh, 1994]** Zadeh L.A. Fuzzy Logic, Neural Network and Soft Computing// Communications of the ACM. – 1994. – Vol.37, №3.
- [Zadeh, 1997]** Zadeh L.A. Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic// Fuzzy Sets and Systems. – 1997. – Vol.90.

УДК 004.8

КОРРЕКТИРОВКА ГИПОТЕЗ ПРОГНОЗА ДЛЯ ИЗВЛЕЧЕНИЯ ЗНАНИЙ О ВРЕМЕННЫХ РЯДАХ¹

И.А. Тимина (*i.timina@ulstu.ru*)

Ульяновский государственный технический университет, Ульяновск

В данной работе предложен подход к решению задач анализа и прогнозирования временных рядов (ВР), входящий в интеллектуальное исследование ВР на основе моделей нечетких временных рядов (НВР), моделей нечетких тенденций (НТ), связанное с корректировкой гипотез прогноза и выделенной с помощью кластеризации ВР доминирующей НТ. На основе результатов анализа рядов пополняется база знаний о ВР новыми правилами.

Введение

В настоящее время информация для анализа различных электронных архивов данных наиболее часто представляется в виде ВР, хранимых в большом количестве в базах данных. Зачастую эти ряды являются нерегулярными, короткими и гетерогенными. Решение задач анализа и прогнозирования одномерных ВР нашло отражение в ряде работ, например, в [Бокс и др., 1974], [Кендэл, 1981], [Zadeh, 1965], [Song, 1993].

Прогнозирование является одной из важных задач в области анализа баз данных ВР. Результаты прогнозирования ВР и тенденций, извлеченных из баз данных полезны во многих сферах деятельности.

В данной работе предлагается алгоритм, корректирующий значения гипотез прогноза. Вводится явная гипотеза от эксперта прикладной области, такая, что НТ ВР Z – это предиктор НТ ВР Y . Понятие «предиктор» используется для обозначения значимых, обеспечивающих наиболее точный прогноз какого-либо явления.

Исходными данными для анализа в проведенных экспериментах являются ВР статистики, собранной в период с 1970-х по 2000-е годы по РСФСР и России в областях: экономика, производство, социальная сфера, здравоохранение, культура.

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 14-01-31092).

Исходные данные представлены четкими значениями и преобразуются в НВР, НТ, прогнозы и рекомендации. Имеются ВР, изменения которых характеризуются большой степенью схождения. Это говорит и о схожем характере поведений локальных тенденций ВР за определенный период. Проводя анализ и прогнозирование с целью выявления выбора наилучшей гипотезы прогноза, можно выявлять информацию о зависимости исследуемого ВР Y от основной его тенденции и от его предиктора ВР Z , сформировать новые правила.

1. Интеллектуальное исследование ВР

Пусть задан дискретный ВР $Y = \{x_i, x_i\} (i=1,2,\dots,n)$. Согласно базовым положениям теории НВР, разработанным *Zadeh* [Zadeh, 1965], *Song* и *Chissom* [Song, 1993], любой конечный дискретный ВР – числовой, нечисловой и смешанный – может быть преобразован в НВР $\tilde{Y} = \{\tilde{x}_i, \tilde{x}_i\} (i=1,2,\dots,n)$, при условии, что его множество значений $X = \{x_i\}$ будет покрыто специальными функциями (нечеткими множествами) $\tilde{x} \in \tilde{X}_i, j=1,2,\dots,m, m < n$ [Ярушкина и др., 2010а].

НТ τ , определяемая на отрезке $[t_i, t_j], t_j \geq t_i$ по значениям \tilde{x}_i, \tilde{x}_j НВР \tilde{Y} , есть нечеткий терм, задающий нечеткое приращение $\tau = \tau((t_i, \tilde{x}_i), (t_j, \tilde{x}_j))$.

Обобщенную модель НВР НТ представим:

$$\tau_i = f(\tau_{i-1}, \dots, \tau_{i-d}),$$

где d – фиксированное число, параметр модели; $\tau_i, \tau_{i-1}, \dots, \tau_{i-d}$ – последовательность НТ; f – некоторая нечеткая зависимость [Афанасьева и др., 2011].

Более детальное описание НТ, носителем которой выступает НВР, определим лингвистическими переменными «НВР», «НТ», «Тип_Тенденции», «Интенсивность_Тенденции», «Длительность_Тенденции» с базовыми конечными терм-множествами $\tilde{X}, \tilde{\mathfrak{Z}}, \tilde{V}, \tilde{A}, \Delta T$, соответственно.

Каждая НТ $\tau \in \mathfrak{Z}$ НВР $\tilde{Y} = \tilde{x}_t, t=1,2,\dots$ может быть представлена структурной моделью в виде кортежа отношения, построенного на декартовом произведении свойств НТ $\tilde{V} \times \tilde{A} \times \Delta T \rightarrow \mathfrak{Z}$:

$$\tau = \langle \tilde{v}, \tilde{a}, \Delta t, \mu \rangle,$$

где τ – наименование НТ из множества \mathfrak{Z} , $\tau \in \mathfrak{Z}$;
 \tilde{v} – тип НТ, $\tilde{v} \in \tilde{V}$, {Падение, Рост, Стабильность};
 $\tilde{\alpha}$ – интенсивность НТ, $\tilde{\alpha} \in \tilde{A}$, {Сильное, Среднее, Слабое};
 Δt – продолжительность НТ, $\Delta t \in \Delta T$;
 μ – функция принадлежности участка НВР, ограниченного интервалом Δt НТ τ .

Классифицируем НТ НВР в зависимости от продолжительности на элементарную $T \in \mathfrak{Z}$ ($\Delta t = 1$), локальную $NT \in N\mathfrak{Z}$ ($1 < \Delta t < n - 1$) и основную (общую) $GT \in G\mathfrak{Z}$ ($\Delta t = n - 1$).

ВР нечетких элементарных тенденций (НЭТ) представим в виде (здесь $TTend$, $RTend$ – процедуры получения нечетких множеств \tilde{v}_t и $\tilde{\alpha}_t$):

$$\tilde{v}_t = TTend(\tilde{x}_t, \tilde{x}_{t+1}), \tilde{\alpha}_t = RTend(\tilde{x}_t, \tilde{x}_{t+1}), \mu_t = \min(\mu(\tilde{x}_t), \mu(\tilde{x}_{t+1})).$$

Согласно предлагаемому в настоящей работе подходу результаты прогнозирования НЭТ необходимо корректировать с учетом основной НТ G_τ . Для идентификации основной НТ G_τ и определения ее компонент по ВР предложен эвристический алгоритм, в котором используемые коэффициенты определены экспериментально. Для функционирования алгоритма требуется, чтобы исходный ВР был преобразован в НВР [Новак и др., 2008].

2. Гипотезы прогнозирования

Для прогнозирования числовых значений и НТ состояния организации в данном продукте используем метод моделирования НЭТ, алгоритм которого заключается в следующем:

1. Построение F-компонент исследуемых ВР. Предлагаемый подход, основанный на методе F-преобразования, позволяет производить декомпозицию и эффективное сглаживание ВР. Причиной применения является то, что коэффициенты F¹-преобразования характеризуют локальные тенденции ряда.

2. Учет выполнения условия эксперта, что взятые для анализа ВР коррелируют.

3. На основе полученных значений коэффициента корреляции используются следующие гипотезы:

1.1. *Гипотеза сохранения тенденции.* Рассматриваем тенденцию за предыдущий месяц. Прогнозируемое значение вычисляется по формуле:

$$Y(t+1) = Y(t) + Tp,$$

где $Y(t+1)$ – прогноз на следующий период времени;

$Y(t)$ – реальное значение в момент времени t ;

Tp – значение тенденции за предыдущий период времени.

1.2. *Гипотеза устойчивости тенденции.* Используем для прогнозирования скользящее среднее

$$Y(t+1) = (1/(T+1)) * [Y(t) + Y(t-1) + \dots + Y(t-T)],$$

где T – количество значений по времени.

Рассматриваем тенденции за предыдущий выбранный период, выбираем преобладающий кластер тенденции и вычисляем прогноз по приведённой выше формуле, строим тенденцию. При одинаковых количествах вхождений тенденций используется оптимистичный прогноз, то есть выбирается наибольшее среднее значение тенденций.

Для нечеткой кластеризации тенденций используем FCM-алгоритм, основанный на минимизации целевой функции:

$$J = \sum_{j=1}^C \sum_{i=1}^N \mu_{ij}^m \|x_i - c_j\|^p,$$

где N – количество показателей;

C – количество кластеров;

μ_{ij} – степень принадлежности объекта i кластеру j ;

m – любое действительное число, большее 1;

x_i – i -й объект набора объектов;

c_j – j -й кластер набора кластеров;

$\|x_i - c_j\|$ – норма, характеризующая расстояние от центра кластера j до объекта i .

Объектами кластеризации является множество элементов ВР $A = \{a_1, \dots, a_n\}$. Задачей FCM-алгоритма является разбиение этого набора на заданное количество кластеров. Вводится конечное множество признаков (тенденций) $P = \{p_1, \dots, p_q\}$, каждый из которых количественно представляет некоторое свойство или характеристику элементов рассматриваемой проблемной области. При этом n есть общее количество объектов данных, а q – общее количество измеримых признаков.

Далее для каждого из объектов кластеризации должны быть измерены все признаки множества P в некоторой количественной

шкале. Тем самым каждому из элементов $a_i \in A$ поставляется в соответствие некоторый вектор $x_i = (x_1^i, x_2^i, \dots, x_q^i)$, где x_j^i – количественное значение признака $p_j \in P$ для объекта $a_i \in A$. Все x_j^i принимают действительные значения. Эти количественные значения признаков тенденций и являются векторами признаков объектов кластеризации. Векторы значений признаков $x_i = (x_1^i, x_2^i, \dots, x_q^i)$ удобно представлять в виде матрицы данных D размерности $n \times q$, каждая строка которой равна значению вектора x_i .

1.3. Прогнозирование на заданный период на основе НЭТ. Алгоритм прогнозирования на период на основе тенденций работает следующим образом. Мы рассматриваем тенденции за предыдущий период за полгода и запоминаем набор тенденций A . Далее ищем такой же набор A тенденций во всех остальных предыдущих полугодиях. Найдя такой набор B , смотрим, какая тенденция C стоит после этого найденного набора B . Строим прогноз, равный тенденции C . Если не существует такого набора B , который совпадал бы с набором A , мы повторяем поиск набора, но уже ищем не полное его совпадение, а с отклонением в одну тенденцию. Так повторяем до тех пор, пока не найдем подходящий набор тенденций B [Тимина, 2013].

3. Алгоритм коррекции гипотез прогноза

Рассмотрим алгоритм корректировки прогнозирования ВР $Y = \{x_i, x_i\}$ ($i = 1, 2, \dots, n$) при условии существования правдоподобности экспертной гипотезы, что НТ ВР $Z = \{z_i, z_i\}$ ($i = 1, 2, \dots, k$) является предиктором ВР Y . Алгоритм 1 включает три шага:

1. Прогнозируются НЭТ ВР Y : $\tau_{t+1}^Y = f(\tau_t^Y)$, где τ_{t+1}^Y – прогнозная НЭТ ВР Y , τ_t^Y – текущая НЭТ ВР Y , f – зависимость в НЭТ ВР Y [9].

2. Прогнозная НЭТ ВР Y регулируется с учетом компонент основных тенденций исследуемого ВР G_{τ_Y} и ВР предиктора. G_{τ_Z} соответственно: $\hat{\tau}_{t+1}^Y = r(\tau_{t+1}^Y, G_{\tau_Y}, G_{\tau_Z})$, где τ_{t+1}^Y – прогнозная НЭТ ВР Y , $\hat{\tau}_{t+1}^Y$ – прогнозная НЭТ ВР Y после корректировки, G_{τ_Y} – основная НТ ВР Y , G_{τ_Z} – основная НТ ВР Z , r – правила корректировки.

3. Используется для получения оценки прогнозного значения числового ВР Y .

На этой основе предлагается следующий алгоритм 2 прогнозирования краткосрочных НТ ВР.

Шаг 1. Преобразование числового ВР в НВР.

Шаг 2. Преобразование НВР в НВР НЭТ: $\tau_t^Y = \langle \tilde{v}_t, \tilde{a}_t, \mu_t \rangle$.

Шаг 3. Построение модели изменения компонент НЭТ ВР Y и их прогнозирование на один период:

$$\tilde{v}_{t+1} = \tilde{v}_t \times \tilde{v}_{t-1} \times \dots \times \tilde{v}_{t-p} \circ R_{\tilde{v}}(t, t-p), \quad \tilde{a}_{t+1} = \tilde{a}_t \times \tilde{a}_{t-1} \times \dots \times \tilde{a}_{t-p} \circ R_{\tilde{a}}(t, t-p).$$

Шаг 4. Прогнозирование числового ВР Y с предварительной дефаззификацией компонент НТ $\hat{\tau}_{t+1}^Y = \langle \tilde{v}_{t+1}, \tilde{a}_{t+1}, \mu_{t+1} \rangle$, $x_{t+1} = x_t + v_{t+1} + a_{t+1}$.

Шаг 5. Применение алгоритма идентификации основной тенденции (см. пункт 3. Алгоритм 1) для ВР Y и вычисление ее компонент $G_{\tau Y} = \langle \tilde{v}_{G_{\tau}}^Y, \tilde{a}_{G_{\tau}}^Y, \mu_{G_{\tau}}^Y \rangle$. Дефаззификация компонент основной НТ ВР Y .

Шаг 6. Шаг 5 для ВР Z .

Шаг 7. Регулировка прогнозной НЭТ ВР Y :

$$\hat{\tau}_{t+1}^Y = r(\tau_{t+1}^Y, G_{\tau Y}, G_{\tau Z}), \quad \hat{\tau}_{t+1}^Y = v_{t+1} \cdot a_{t+1} + v_{G_{\tau}}^Y \cdot a_{G_{\tau}}^Y + v_{G_{\tau}}^Z \cdot a_{G_{\tau}}^Z.$$

Шаг 8. Вычисление отрегулированного прогнозного значения числового ВР Y на один период: $x_{t+1}^i = x_t + \hat{\tau}_{t+1}^Y$.

Для проверки качества полученных значений прогноза использовалась следующая градация оценок (табл. 1):

Таблица 1

Оценки качества

Квадрат среднеквадратичной ошибки	$MSE = \frac{1}{n} \sum_{t=1}^n (x_t - x_t^i)^2$
Средняя квадратичная ошибка	$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (x_t - x_t^i)^2}$

4. Формирование правил

Описанные предыдущие решения задач анализа предлагается заложить в основу инструментария для генерации правил, которые впоследствии будут дополнительно охарактеризованы пользователем. Так, полученные результаты на основе корректировок гипотез прогноза с учетом наличия минимальной ошибки MSE предлагается использовать для генерации правил следующих типов:

1. Если минимальную ошибку имеет гипотеза без корректировки, то ВР Y считать не зависимым от общей тенденции ВР Y и от ВР Z .

2. Если ошибка минимальна у гипотезы с корректировкой по основной тенденции ВР Y , то ВР Y понимать как не зависимым от ВР Z , но зависимым от общей тенденции самого ВР Y .

3. Если гипотеза откорректирована по ВР Z и ее ошибка минимальна, то ВР Y рассматривать зависимым и от основной тенденции ВР Y , и от ВР Z .

Данная конструкция составления правил позволит осуществить гибкий подход к анализу электронного архива данных, представленных ВР, пользователей. Производя такое формирование, пользователям в абстрактной форме будет предложено охарактеризовать возникающие распределения гипотез прогноза. Сведения в дальнейшем лягут в основу базы знаний.

5. Результаты вычислительного эксперимента

Для анализа в проведенных экспериментах использовались ВР статистики, собранной в период с 1970-х по 2000-е годы по РСФСР и России в областях: экономика, производство, социальная сфера, здравоохранение, культура. ВР взяты с условием наличия высокой корреляции.

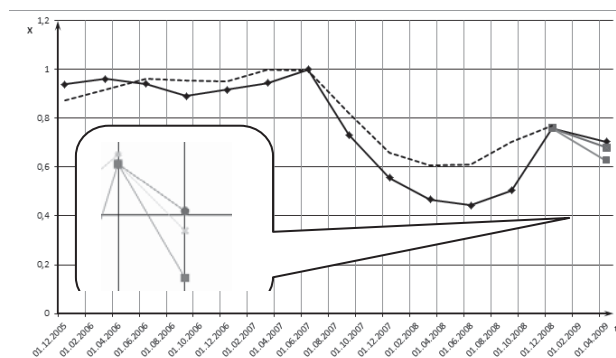


Рис 1. Гипотеза прогнозирования ВР Y на период на основе НЭТ с учетом корректировки

На рисунке 1 представлен вариант прогнозирования с использованием гипотезы, где ось x – значение ВР, ось t – время; «◆◆◆»

обозначает ВР Y ; «—●—» обозначает ВР Z ; «-----» обозначает прогноз без учета корректировки; «—*—» обозначает прогноз τ_{t+1}^Y после корректировки ВР Y по основной тенденции ВР Y ; «—■—» обозначает прогноз $\hat{\tau}_{t+1}^Y$ после корректировки ВР Y по основной тенденции ВР Z .

В таблицах 2 и 3 показаны полученные результаты оценок прогнозирования и анализа ВР.

Таблица 2

Оценки прогнозирования

Оценка	Гипотеза сохранения НЭТ		Гипотеза устойчивости НЭТ		Гипотеза на основе НЭТ	
	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
<i>MSD</i>	0,0014	0,0001	0,2778	0,3285	0,0592	0,0389
<i>RMSD</i>	0,0373	0,0088	0,5271	0,5731	0,2433	0,1973

Таблица 3

Результаты анализа

Ряд	Общая тенденция	Корреляция	Гипотеза сохранения НЭТ	Гипотеза устойчивости НЭТ	Гипотеза на основе НЭТ
НВР Y	Падение среднее	0,98 (сильная)	Рост средний	Рост средний	Падение среднее
НВР Z	Рост малый				

Поскольку гипотеза сохранения тенденции не учитывает основную тенденцию ряда, результативнее использовать гипотезу на основе НЭТ.

Согласно предложенному алгоритму 2 на первой фазе при прогнозировании ВР Y (рис. 1) получили прогнозную НЭТ «Падение малое» и ошибки прогнозных значений ВР ($MSE=0,0592$, $RMSE=0,2433$), затем, на второй фазе применяя алгоритм 1, определили основную НТ ВР Y «Падение среднее» и основную НТ ВР Z «Рост малый», дефазифицированные значения которых были использованы для коррекции прогнозных значений ВР Y полученного на первой фазе ($MSE=0,0389$, $RMSE=0,1973$).

В результате формируется база правил, с помощью которых облегчается работа пользователя в плане прогнозирования, резюмирования и анализа исходных ВР.

СПИСОК ЛИТЕРАТУРЫ

- [**Бокс и др., 1974**] Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Пер. с англ. и под ред. В. Ф. Писаренко. М., Мир. 1974.
- [**Кендэл, 1981**] Кендэл М. Временные ряды. Пер. с англ. и пред. Ю. П. Лукашина. М., Финансы и статистика. 1981.
- [**Zadeh, 1965**] Zadeh A. Lotfi. Fuzzy Sets // Information and Control. 1965.
- [**Song, 1993**] Song Q., Chissom B. Fuzzy time series and its models // Fuzzy Sets and Systems. 1993. № 54.
- [**Ярушкина и др., 2010а**] Ярушкина Н.Г., Афанасьева Т.В., Перфильева И.Г. Интеллектуальный анализ временных рядов: учебное пособие. – Ульяновск: УлГТУ. 2010.
- [**Афанасьева и др., 2011**] Афанасьева Т.В., Ярушкина Н. Г. Нечеткий динамический процесс с нечеткими тенденциями в анализе временных рядов // Вестник Ростовского государственного университета путей сообщения. 2011. №3.
- [**Новак и др., 2008**] Интегральный метод принятия решений и анализа нечетких временных рядов / В. Новак [и др.] // Программные продукты и системы. 2008. № 4.
- [**Тимина, 2013**] Тимина И.А. Нечеткая зависимость как метод решения задач интеллектуального анализа временных рядов // Автоматизация процессов управления. Ульяновск: НПО «Марс». 2013. № 3(33).
- [**Ярушкина и др., 2010б**] Ярушкина Н.Г., Афанасьева Т.В. Метод нечеткого моделирования и анализа тенденций временных рядов // Интеллектуальные системы управления. Коллективная монография / под ред. акад. РАН С.Н. Васильева. – М.: Машиностроение, 2010.
- [**Ярушкина и др., 2007**] Ярушкина Н.Г., Афанасьева Т.В. Нечеткие временные ряды как инструмент для оценки и измерения динамики процессов // Датчики и системы. 2007. № 12.
- [**Ярушкина и др., 2010с**] Ярушкина Н.Г., Перфильева И.Г., Афанасьева Т.В. Интегральный метод нечеткого моделирования и анализа нечетких тенденций // Автоматизация процессов управления. 2010. № 2 (20).

УДК 004.85

НАСТРОЙКА ПАРАМЕТРОВ НЕЧЕТКОГО КЛАССИФИКАТОРА АЛГОРИТМОМ «КУКУШКИН ПОИСК»¹

И.А. Ходашинский (*hodashn@rambler.ru*)

Д.Ю. Минина (*pound_2007@mail.ru*)

Томский государственный университет систем управления
и радиоэлектроники, г. Томск, Россия

В работе представлен алгоритм «кукушкин поиск» для настройки параметров нечеткого классификатора. Описаны результаты экспериментов на четырех наборах данных из репозитория KEEL: balance, pima, vura, iris. Даны сравнения с аналогами.

Введение

Разделение множества объектов по заданным признакам на группы или классы – проблема, существующая во многих областях человеческой деятельности. Результатом классификации является метка класса, к которому принадлежит объект классификации. Классификатор – это алгоритм, выполняющий отображение из области признаков в область меток класса. Существует множество различных типов классификаторов, которые используют различные подходы, чтобы выполнить указанное отображение: линейный дискриминантный анализ; байесовская классификация, деревья решений; нейронные сети; метод опорных векторов, нечеткие системы.

Популярность классификаторов с нечеткими правилами обусловлена тем, что представление знаний в виде правил интуитивно понятно человеку, а также применением лингвистического описания вместо числового, когда нет необходимости в указании точных значений входных признаков. В то же время использование принципа «победитель получает все» позволяет для нового классифицируемого объекта определить класс, к которому этот объект наиболее близок.

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 12-07-00055а).

Настройка классификатора заключается в нахождении параметров классификатора, на которых заранее сформированное множество объектов (обучающая выборка) удовлетворяет заданному качеству классификации.

Настройка параметров нечеткого классификатора – это оптимизация параметров функций принадлежности, для этих целей наряду с алгоритмами оптимизации, основанными на производных, применяются метаэвристики, чаще всего эволюционные алгоритмы [Nettgera, 2008].

Цель предлагаемой работы – описание метода настройки параметров нечеткого классификатора с помощью популяционного алгоритма «кукушкин поиск».

Выбор указанной метаэвристики обусловлен проведенными в работе [Yang et al., 2009] сравнениями с генетическим алгоритмом и алгоритмом роящихся частиц. Эксперименты показали, что алгоритм «кукушкин поиск» показывает лучшие результаты при решении задач оптимизации по сравнению с указанными алгоритмами.

1. Постановка задачи

Основная идея нечеткого классификатора состоит в описании предполагаемого кластера нечетким прототипом, размерность которого определена размерностью пространства исследуемых данных. Таким образом, i -й кластер определяется нечетким правилом следующего вида:

$$R_{ij} : \text{ЕСЛИ } x_1=A_{1i} \text{ И } x_2=A_{2i} \text{ И } x_3=A_{3i} \text{ И } \dots \text{ И } x_n=A_{ni} \text{ ТО class}=c_j,$$

где $\mathbf{x} = (x_1, x_2, x_3, \dots, x_n)$ – вектор признаков классифицируемого объекта;

A_{ki} – нечеткий терм, характеризующий k -ый признак в i -том правиле ($i \in [1, R]$), R – число правил; c_j – идентификатор j -того класса, $j \in [1, m]$.

Нечеткая классификация описывается функцией

$$f : \mathfrak{R}^n \rightarrow [0,1]^m,$$

которая относит классифицируемый объект к каждому классу с определенной степенью принадлежности, вычисленной следующим образом:

$$\beta_j(\mathbf{x}) = \sum_{R_j} \prod_{k=1}^n A_{ki}(x_k), j = 1, 2, \dots, m.$$

Выходом классификатора является класс, определяемый следующим образом:

$$\text{class} = c_{j^*}, \quad j^* = \arg \max_{1 \leq j \leq m} \beta_j.$$

Нечеткий классификатор может быть представлен как функция

$$c = f(\mathbf{x}, \boldsymbol{\theta}),$$

где $\boldsymbol{\theta}$ – вектор, описывающий базу правил.

Пусть дано множество обучающих данных (таблица наблюдений) $\{(\mathbf{x}_p; c_p), p = 1, \dots, z\}$, определим следующую единичную функцию

$$\text{delta}(p, \boldsymbol{\theta}) = \begin{cases} 1, & \text{если } c_p = f(\mathbf{x}_p, \boldsymbol{\theta}) \\ 0, & \text{иначе} \end{cases}, \quad p = 1, 2, \dots, z,$$

тогда численный критерий качества классификации может быть выражен следующим образом:

$$E(\boldsymbol{\theta}) = \frac{\sum_{p=1}^z \text{delta}(p, \boldsymbol{\theta})}{z}.$$

Проблема настройки сводится к проблеме поиска максимума заданной функции в многомерном пространстве, координаты которого соответствуют параметрам нечеткого классификатора. Для оптимизации $\boldsymbol{\theta}$ предлагается использовать алгоритм «кукушкин поиск».

2. Алгоритм «кукушкин поиск»

Оптимизационный алгоритм «кукушкин поиск» (АКП) – мета-эвристика, построенная на основе имитации способа размножения кукушки, когда она находит недавно построенные гнезда и подкладывает в них свои яйца (заменяет своими), которые в итоге могут быть выкинуты хозяином гнезда [Yang et al., 2010]. В основе алгоритма лежат три следующих правила:

1) кукушка откладывает по одному яйцу в случайно выбранное гнездо, которое представляет собой решение;

2) часть лучших решений будет перенесена в следующее поколение;

3) количество гнезд фиксировано и есть вероятность того, что хозяин может обнаружить чужое яйцо; в этом случае хозяин может

выбросить яйцо из гнезда или вовсе отказаться от гнезда и построит новое на новом месте.

Важной составляющей АКП является использование полетов Леви для локального и глобального поиска [Mantegna, 1994]. Процесс полета Леви является случайным блужданием, которое характеризуется серией скачков, обусловленных функцией плотности вероятности с «толстыми» хвостами, за счет которых вероятность значительных отклонений от среднего больше, чем у нормального распределения.

Собственно алгоритм представлен ниже.

Шаг 1. Инициализация исходной популяции.

Задается S – размер популяции $\Theta = (\theta^s, s \in [1, S])$. Случайным образом задается «начальное положение кукушки», являющееся текущим решением θ^{cur} . При этом каждый нечеткий терм (здесь и в дальнейшем) в каждом векторе θ^s и в векторе θ^{cur} задается в границах диапазона с учетом того, что левая граница каждого последующего терма отдельной переменной должна находиться правее левой границы предыдущего терма. Задается p – вероятность, с которой гнездо может быть «покинуто» хозяином, т.е. вероятность удаления векторов из множества Θ . Задается количество итераций N в качестве критерия остановки. Количество генерируемых векторов l полагаем равным S .

Шаг 2. Случайным образом генерируется l векторов θ .

Шаг 3. Генерация нового решения на основе полетов Леви.

Выполняется «случайное перемещение кукушки», которое выражено изменением θ^{cur} по закону Леви

$$\theta^{cur} = \theta^{cur} + Levi,$$

где $Levi$ – случайный прыжок полета Леви, вычисляемый по правилу

$$Levi = \gamma \cdot u / |v|^{1/\beta},$$

где γ – коэффициент прыжка полета Леви; параметр β принимает значения из интервала $[1, 2]$; u, v – нормально распределенные величины $v \sim N(0; \sigma_v^2)$, $u \sim N(0; \sigma_u^2)$,

$$\sigma_v^2 = 1, \sigma_u^2 = \left\{ \frac{\Gamma(1 + \beta) \sin(\pi\beta / 2)}{\Gamma[(1 + \beta) / 2] \cdot 2^{(\beta-1)/2}} \right\}^{1/\beta},$$

$\Gamma(x)$ – гамма-функция.

Случайным образом выбирается другое решение θ^s из популяции Θ .

Шаг 4. Оценка качества решения.

Сравниваются значения функций $E(\theta^s)$.

Если $E(\theta^{cur}) < E(\theta^s)$, то $\theta^{cur} = \theta^s$, иначе $\theta^s = \theta^{cur}$.

Шаг 5. Удаление «неудачных гнезд» (решений).

Если не выполнено заданное количество итераций N ,

то выбирается заранее заданное количество t «худших» решений (с наименьшим значением E); для каждого из них генерируется случайное число h в диапазоне $[0, 1]$, и если значение h для решения оказывается больше значения заданной вероятности p , гнездо-решение удаляется; вместо удаленного решения формируется новое: l полагается равным количеству уничтоженных векторов-решений, переход на шаг 2,

иначе, выбирается лучшее решение (с максимальным значением $E(\theta)$); конец работы алгоритма

3. Эксперимент

Для оценки эффективности нечеткого классификатора, настроенного приведенным алгоритмом, были проведены тесты на четырех наборах данных из репозитория KEEL (<http://www.keel.es>): balance, rima, vira, iris по схеме кросс-валидации. В этой схеме набор разбивался на пять файлов, 80% набора помещалось в обучающую выборку, а остальные 20% в тестовую. Параметры алгоритма и нечеткого классификатора следующие: число итераций – 5000; количество особей в популяции – 10; количество уничтожаемых худших решений – 3; коэффициент прыжка полета Леви γ – 0,1; параметр $\beta = 1,1$; функции принадлежности – треугольные. Число правил в классификаторах на наборах данных: balance – 3, rima – 2, vira – 2, iris – 81.

Результаты проведенных экспериментов были сопоставлены с результатами из работы [Alcala-Fdez et al., 2011], где испытания проводились на тех же данных. В таблицах 1–4 в столбце «среднее» указаны усредненные значения процента правильной классификации, в столбце «СКО» – среднеквадратическое отклонение.

Полученные классификаторы имеют хорошие способности к обучению (процент правильной классификации на обучающей выборке) и не менее хорошие прогностические способности (процент правильной классификации на тестовой выборке).

Таблица 1

Набор данных balance

Алгоритм	Обучающая выборка		Тестовая выборка	
	среднее	СКО	среднее	СКО
Ant Miner	73,65	3,38	70,24	6,21
CORE	68,64	2,57	70,08	7,11
HIDER	75,86	0,40	69,60	3,77
SGERD	76,96	2,27	75,19	6,27
TARGET	77,29	1,57	75,62	7,27
Наш алгоритм	89,88	2,17	87,74	3,28

Таблица 2

Набор данных rima

Алгоритм	Обучающая выборка		Тестовая выборка	
	среднее	СКО	среднее	СКО
Ant Miner	71,86	2,84	66,28	4,26
CORE	72,66	2,62	73,06	6,03
HIDER	77,82	1,16	73,18	6,19
SGERD	73,71	0,40	73,71	3,61
TARGET	73,42	2,62	73,02	6,61
Наш алгоритм	76,25	1,30	70,87	1,43

Таблица 3

Набор данных vura

Алгоритм	Обучающая выборка		Тестовая выборка	
	среднее	СКО	среднее	СКО
Ant Miner	80,38	3,25	57,25	7,71
CORE	61,93	0,89	61,97	4,77
HIDER	73,37	2,70	65,83	10,04
SGERD	59,13	0,68	57,89	3,41
TARGET	68,86	0,89	65,97	1,41
Наш алгоритм	71,80	1,39	68,48	4,83

Таблица 4

Набор данных iris

Алгоритм	Обучающая выборка		Тестовая выборка	
	среднее	СКО	среднее	СКО
Ant Miner	97,26	0,74	96,00	3,27
CORE	95,48	1,42	92,67	4,67
HIDER	97,48	0,36	96,67	3,33
SGERD	97,33	0,36	96,67	3,33
TARGET	93,50	2,42	92,93	4,33
Наш алгоритм	98,63	0,58	96,67	2,87

Заключение

В работе представлена процедура построения нечетких классификаторов на основе алгоритма «кукушкин поиск». Работоспособность нечетких классификаторов, настроенных приведенным алгоритмом, проверена на четырех наборах данных из репозитория KEEL. Приведенные сравнения разработанного алгоритма с аналогами показали его эффективность при решении задачи оптимизации параметров нечетких классификаторов.

СПИСОК ЛИТЕРАТУРЫ

[Alcala-Fdez et al., 2011] Alcala-Fdez J., Fernandez A., Luengo J., Derrac J., Garcia S., Sanchez L., Herrera F. KEEL Data-mining software Tool: data, set repository, integration of algorithms and experimental analysis framework // J. of Mult.-Valued Logic and Soft Computing. 2011. Vol. 17.

[Herrera, 2008] Herrera F. Genetic fuzzy systems: taxonomy, current research trends and prospects // Evolutionary Intelligence. 2008. V. 1.

[Mantegna, 1994] Mantegna R.N. Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes // Physical Review E. 1994. V.49.

[Yang et al., 2009] Yang X.-S., Deb S. Cuckoo Search via Lévy flights // Proc. World Congr. on Nature & Biologically Inspired Computing. (NaBIC 2009), India. 2009.

[Yang et al., 2010] Yang X.-S., Deb S. Engineering optimisation by cuckoo search // Int. J. Mathematical Modelling and Numerical Optimisation. 2010. V. 1.

УДК 004.82

**СЕМАНТИЧЕСКИЕ ТЕХНОЛОГИИ ГЕНЕРАЦИИ
УЧЕБНЫХ ВОПРОСОВ¹****М.М. Аюпов** (*madehur@mail.ru*)**О.А. Невзорова** (*onevzoro@gmail.com*)**Н.А. Прокопьев** (*nikolaskuji@gmail.com*)**Д.Ш. Сулейманов** (*dvdt.slt@gmail.com*)НИИ «Прикладная семиотика АН РТ», Казанский (Приволжский)
федеральный университет, г. Казань, Россия

В статье рассматривается модель автоматической генерации вопросов на основе форм всевозможных запросов к реляционной базе данных, в таблицах которой представлена модель знаний предметной области «Информатика». Каждый учебный вопрос описывается на трех формальных уровнях: онтологическом, лексическом и уровне реализации (SQL-форма запроса). Анализируются результаты экспериментов по генерации учебных вопросов для школьного курса информатики.

Введение

Как показывает практика, тестирующая подпрограмма является важным компонентом современных автоматизированных обучающих систем. При этом наиболее наукоемкой и ресурсоемкой задачей является разработка качественного теста, позволяющего достоверно оценить уровень знаний обучаемого. Соответственно, актуальным является разработка программных средств, обеспечи-

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 12-07-00550-а).

вающих автоматическую генерацию вопросов и моделей ответов, что может способствовать уменьшению трудоемкости подготовки тестовых заданий и повышению качества контроля знаний обучаемых. Автоматическая генерация вопроса по учебному тексту является сложной задачей и в общем случае требует полнофункционального анализа текста, включающего морфологический, синтаксический и семантический анализы.

Рассмотрим основные методы, используемые в задаче автоматической генерации вопросов. В [Naveed et al., 2013; Papasalouros et al., 2008] рассматривается подход к генерации вопросов с вариантами ответов, который заключается в распознавании наиболее важных термов и именованных сущностей и определении семантических отношений между ними без конкретизации типов этих отношений или специфики их языковой реализации. Подход в [Weiming et al., 2008] использует генерацию вопросов на основе шаблонов, выделяемых из коллекции медицинских статей. В [Tianyong et al., 2008] используется метод автоматической генерации шаблонов-вопросов, с дальнейшим использованием системы WordNet и специальной онтологии тегов для разметки объектов и связей и порождения иерархии входных фактов. В [Tahani et al., 2012] описывается подход к генерации вопросов с ответами множественного (выборочного) типа с использованием онтологий. В [Arjun et al., 2013] представлен метод порождения вопросов с вариантами ответа, который использует вопросы с вариантами ответов, находящихся в публичном доступе. В [Кручинин, 2004] дан алгоритм генерации вопросов с использованием механизма генерации комбинаторных множеств на основе деревьев И/ИЛИ. В [Кручинин и др., 2004] рассматривается метод автоматической генерации вопросов с вариантами ответов с использованием шаблонов на основе табличного представления данных.

В статье рассматриваются модель генератора учебных вопросов, использующего базу знаний онтологического типа, представленную в структурах реляционной базы данных, а также многоуровневое представление модели вопроса. Структура базы знаний определяет основные характеристики генератора:

- количество типов вопросов, которое может быть получено данным генератором,
- сложность полученных вопросов,

– корректность генерируемых вопросов.

В статье описывается база знаний, разработанная для школьного курса «Информатика». Для данного курса выполнены эксперименты по генерации учебных вопросов. Можно отметить, что предложенная модель генерации позволяет создавать более сложные вопросы, чем в подходах, описанных в работах [Naveed et al., 2013; Weiming et al., 2008; Tianyong et al., 2008; Tahani et al., 2012; Papasalouros et al., 2008].

1. Модель знаний предметной области «Информатика»

Для построения модели генерации учебных вопросов была выбрана предметная область «Информатика», представленная в учебнике «Информатика» на русском языке для учеников 8 класса средних школ. Данный учебный курс состоит из трех основных разделов и 19 подразделов, включающих базовые знания об архитектуре компьютера, основах процессов кодирования и обработки текстовой и графической информации.

Знания предметной области «Информатика» представляются в таблицах реляционной базы данных (БД). Для представления знаний определены 8 типов таблиц, в которых представляются объекты, бинарные предикаты (бинарные отношения), свойства объектов, бинарные простые функции (функции, определяемые на множестве объектов), бинарные сложные функции (функции, определяемые на множестве функций), n-местные сложные функции, синонимы объектов и классы объектов. Реляционная модель представления онтологии позволяет использовать встроенные средства для генерации сложных SQL-запросов с последующим преобразованием SQL-формы в текстовое представление. Поскольку для представления запросов используются такие сложные представления, как композиция функций (предикаты второго порядка), модели онтологий на основе OWL являются менее пригодными для применения.

Для представления знаний о предметной области в таблицах многоместных отношений разработана архитектура БД и реализован программный модуль для автоматизированного заполнения таблиц БД объектами, выделенными из текста. Главная форма приложения (рис. 1) содержит две основные рабочие части: область для работы с текстом и область для работы с таблицами.

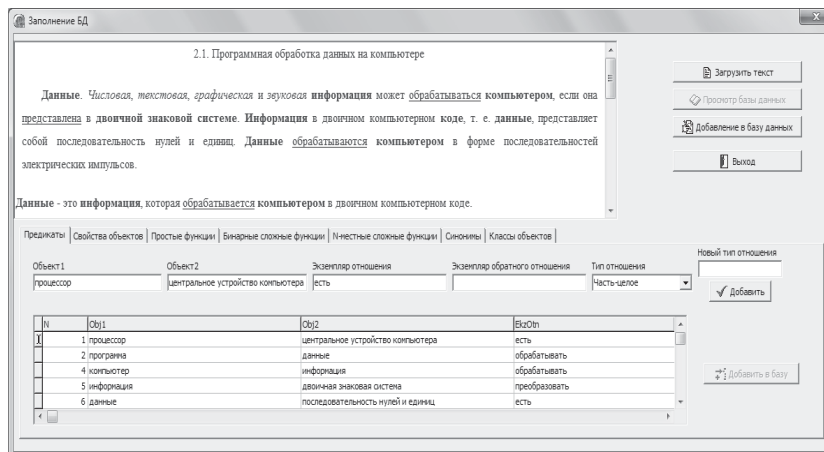


Рис. 1. Главная форма программы формирования модели предметной области

В начале работы в область для работы с текстом загружается учебный текст, в котором автоматически размечаются объекты, отношения и значения свойств объектов из базы данных. Новые экземпляры БД создаются путем ручной разметки и автоматически заносятся в соответствующую таблицу БД. Выбор таблицы в экранной области для работы с таблицами определяет тип контекстного меню, с помощью которого происходит последующее заполнение таблицы данными из текста.

При выборе нужного пункта контекстного меню текстовые объекты формируют конкретный экземпляр объекта или конкретный экземпляр отношения в строке соответствующей таблицы отношений или конкретный экземпляр свойства в таблице свойств и т.д.

При заполнении полей таблицы происходит автоматическая пословная лемматизация текстовых данных. Дальнейшее развитие функционала программы может быть связано с введением новых схем отношений и разработкой методов автоматического распознавания отношений в тексте.

2. Модель генерации учебных вопросов и эксперименты

Вопросы по предметной области генерируются как всевозможные запросы к БД. Нами выделяются простые и сложные запросы. Простыми запросами называются запросы, которые задаются от-

носителем одной таблицы. Если при построении запросов используются несколько таблиц, такие запросы называются сложными. Каждый контрольный вопрос описывается на трех уровнях:

- онтологический уровень (представление запроса в структурах онтологии);
- лексический уровень (представление запроса в естественно-языковой (ЕЯ)-форме);
- уровень реализации (SQL-форма запроса).

Примером простого запроса является запрос типа: «Указать объект X_1 , который состоит в заданном отношении R_1 с заданным объектом Y_1 ». Рассмотрим для этого запроса перечисленные выше три уровня описания.

Онтологический уровень: *Объект1(= X_1)? + Имя отношения (=R₁) + Объект2(= Y_1).*

Лексический уровень: *Объект1(= X_1)? → «что?», Имя отношения (=R₁) → «являются», Объект2 (=Y₁) → «центральное устройство компьютера».*

ЕЯ-форма: «Что является центральным устройством компьютера?».

Уровень реализации: SQL-запрос к соответствующей таблице БД.

К сложным запросам относятся, например, запросы типа:

1. Указать общие значения $V_1...V_M$ заданного свойства P для объектов $X_1...X_N$ из заданного класса K .
2. Указать объекты $X_1...X_N$ из заданного класса K , имеющие заданные значения $V_1...V_M$ заданного свойства P .

В модели выделены 15 типов вопросов. Например, вопрос типа 1 есть вопрос относительно состава заданного объекта. Ответом на вопрос являются номера объектов, являющихся конструктивными частями заданного объекта. Пример: «Из каких частей состоит имя файла?» – Ответ конструируется с указанием двух составляющих – собственно имя файла и расширение файла. Вопрос типа 6 есть вопрос на определение некоторой сущности. Ответом на вопрос является текстовое определение заданной сущности. Пример: «Определите понятие «программа»?» – Возможный ответ: «Программа есть последовательность команд, которую выполняет компьютер в процессе обработки данных».

Для всех типов вопросов полностью реализовано два уровня представления:

– онтологический уровень автоматически строится из структур таблиц БД,

– уровень реализации выполняется с помощью соответствующего SQL- запроса.

Сложность реализации лексического уровня связана с синтезом ЕЯ-формы вопроса. Для решения этой проблемы предлагается использовать подход на основе расширяемых шаблонов, который реализуется с помощью специального модуля генерации вопросов.

Все шаблоны имеют общую структуру следующего вида:

Имя шаблона:

Описание типа вопроса:

Строка SQL-запроса: «SELECT <названия полей> FROM <таблицы>
WHERE <условия>»

Поля ответа:

«<имя поля, формирующего ответ>»,
[«<имя поля, формирующего ответ>»],

...

Тип ответа: «Сущность»/«Функция»/«Значение не определено»

Вопросительные слова:

«<вопросительное слово>»,
[«<вопросительное слово>»],

...

Главный объект вопроса:

Тип главного объекта: «Сущность»/«Функция»

Падеж: «им»/«род»/«дат»/«вин»/«твор»/«пред»/

Число: «ед»/«мн»

Наличие вопросительного знака: true/false

Приведем примеры вопросов, относящихся к выделенным типам, а также соответствующий SQL-запрос, используемый при генерации (табл.1).

Таблица 1

Примеры сгенерированных вопросов

Тип вопроса	Текстовая форма вопроса	SQL-запрос
1	Из каких частей состоит имя файла?	SELECT part_object_id FROM ObjectParts WHERE whole_object_id= ?

Продолжение таблицы 1

6	Определите понятие «программа»?	SELECT entity_definition FROM Entities WHERE entity_id = ? AND entity_definition IS NOT NULL»
12	Опишите процесс программной обработки данных	SELECT element_function_id FROM CompositeFunctions WHERE composite_function_id = ?

Тестирование модуля генерации вопросов выполнялось на учебных вопросах, взятых из учебника по курсу «Информатика» для 8 класса средней школы. Результаты классификации учебных вопросов представлены в табл. 2.

Таблица 2

Классификация учебных вопросов

	Всего вопросов	По типам														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Количество	139	6	7	9	2	8	27	14	7	2	4	4	7	7	5	30
%	100	4	5	6.5	1.5	5.7	19.3	10	5	1.5	3	3	5	5	3.5	22

Классификация вопросов из сгенерированной выборки (по всем данным всех таблиц экспериментальной БД) приведена в табл. 3.

Таблица 3

Классификация всех сгенерированных вопросов

	Всего вопросов	По типам														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Количество	161	6	10	20	6	8	27	21	16	31	4	4	8	0	0	0
%	100	3.7	6.2	12.4	3.7	5	16.8	13	9.9	9.3	2.5	2.5	5	0	0	0

Из выборки сгенерированных вопросов были исключены вопросы, не имеющие смысла. Содержательная выборка вопросов представлена в табл. 4.

Таблица 4

Классификация содержательной выборки сгенерированных вопросов

	Всего вопросов	По типам														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Количество	140	6	10	20	6	8	27	20	7	20	4	4	8	0	0	0
%	100	4.3	7.1	14.3	4.3	5.7	19.3	14.3	5.0	14.3	2.9	2.9	5.7	0	0	0

При сравнении выборки содержательных вопросов с выборкой всех сгенерированных вопросов видно, что в основном несодержательные вопросы принадлежали к тем типам вопросов, которые получены из таблицы «Простые функции». Таким образом, 13% вопросов оказались несодержательными.

Сравнение выборки содержательных вопросов с выборкой вопросов из учебника показано в таблице 5.

Таблица 5

Результаты сравнения выборок вопросов

	Учебная выборка	Объединение выборок	Пересечение выборок	Уникальные вопросы из учебной выборки	Уникальные вопросы из сгенерированной выборки
Количество	139	182	97	42	43
%	100	31	70	30	31

Анализ результатов сравнения показывает, что сгенерированные вопросы покрывают 70% всех вопросов учебника, а база вопросов увеличивается на 31%.

Уникальными для учебной выборки вопросами являются вопросы, которые не относятся ни к одному из введенных типов. Наиболее существенно база вопросов расширяется за счет типов 2, 3, 4, 7, 9. Вопросы таких типов возникают вследствие наличия в базе данных дополнительных связей, описывающих предметную область.

Заключение

В статье предложена модель генерации учебных вопросов, в которой знания о предметной области представляются в таблицах многоместных отношений, которые реализуются в виде БД. Вопросы по учебному курсу генерируются как всевозможные запросы к БД. Разработанная модель открыта для пополнения знаний. Проведенные эксперименты по генерации учебных вопросов показали хорошую точность по сравнению с обучающей выборкой практически для всех определенных типов вопросов. Уникальные вопросы учебной выборки, для которых отсутствует модель генерации в текущей реализации, являются объектом для дополнитель-

ного изучения. Пути улучшения качества генерации могут быть связаны как с расширением атрибутов таблиц БД, так и с проектированием механизмов вывода на данных БД, что составляет содержание новых задач, связанных с генерацией вопросов.

СПИСОК ЛИТЕРАТУРЫ

[Кручинин, 2004] Кручинин В.В. Использование деревьев И/ИЛИ для генерации вопросов и задач // Вестник Томского государственного университета. 2004. № 284. С. 182–186.

[Кручинин, 2004] Кручинин В.В., Морозова Ю.В. Модели генераторов вопросов для компьютерного контроля знаний // Открытое и дистанционное образование. 2004. № 2. С. 52–62.

[Arjun et al., 2013] Arjun Singh Bhatia, Manas Kirti, Sujan Kumar Saha Automatic Generation of Multiple Choice Questions Using Wikipedia. Pattern Recognition and Machine Intelligence // Lecture Notes in Computer Science. Volume 8251, 2013, pp 733–738.

[Naveed et al., 2013] Naveed Afzal, Ruslan Mitkov Automatic generation of multiple choice questions using dependency-based semantic relations // Soft Computing, A Fusion of Foundations, Methodologies and Applications, Springer-Verlag Berlin Heidelberg, 2013.

[Papasalouros et al., 2008] A. Papasalouros, K. Kanaris, K. Kotis Automatic Generation Of Multiple Choice Questions From Domain Ontologies // E-Learning Conference. Amsterdam, The Netherlands, 2008, pp. 427–434.

[Tianyong et al., 2008] Tianyong Hao, Wanpeng Song, Dawei Hu, Wenyin Liu Automatic Generation of Semantic Patterns for User-Interactive Question Answering // Information Retrieval Technology. Lecture Notes in Computer Science. Volume 4993, 2008, pp 632–637.

[Tianyong et al., 2008] Tahani Alsubait, Bijan Parsia, Uli Sattler Automatic generation of analogy questions for student assessment: an Ontology-based approach. Research in Learning Technology, 2012, pp. 95–102.

[Weiming et al., 2008] Weiming Wang, Tianyong Hao, Wenyin Liu Automatic Question Generation for Learning Evaluation in Medicine // Advances in Web Based Learning – ICWL 2007. Lecture Notes in Computer Science. Volume 4823, 2008, pp 242–251.

УДК 004.82

СУБЪЕКТИВНАЯ КЛАССИФИКАЦИЯ ПРЕДМЕТНЫХ ОНТОЛОГИЙ¹

Т.А. Гаврилова (*gavrilova@gsom.pu.ru*)

Высшая школа менеджмента СПбГУ, Санкт-Петербург

В работе рассмотрены основные методологические проблемы формирования онтологий. Предложена новая классификация онтологий, связанная с проблемой оценки по различным критериям. Акцент делается на возможных ошибках при разработке практических предметных онтологий.

Статья посвящена методологическим аспектам формирования практических онтологий для порталов знаний. Портал знаний в широком смысле – это обобщающий термин, который сегодня используется для корпоративных порталов любой конфигурации. Многие разработчики используют его как синоним информационного портала. В узком смысле портал знаний предприятия (Enterprise Knowledge Portal -ЕКР) содержит действительно базу знаний, включая опыт, успешные практики, технологии, бизнес-процессы, ошибки и другие составляющие интеллектуального капитала компании. Портал знаний предприятия – результат эволюции традиционного портала под влиянием теории управления знаниями (УЗ). ЕКР не только предоставляет средства доступа к информации, но и позволяет пользователям взаимодействовать друг с другом, помогая связывать информацию с коллективным пониманием, системой ценностей и опытом (Grammer, 2000). ЕКР дает возможность принимать оптимальные управленческие решения, поскольку сочетает приобретенные знания с данными и бизнес-контентом и служит центром обучения новичков на опыте заслуженных специалистов и экспертов.

Ближе всего к понятию ЕКР приблизились разработчики в хрестоматийном эдинбургском проекте Enterprise Project [Uschold, et al., 1997], который ставил задачу создания платформы для:

1) сбора и спецификации моделей бизнеса,

¹ Работа выполнена при финансовой поддержке РФФИ (проект 14-07-00294).

2) описания стратегии, организационных структур, ресурсов и продуктов,

3) спецификации бизнес-проблем и бизнес-требований,

4) идентификации решений и альтернатив проектирования.

При этом, как и во всех других проектах в данной области [Гурьянова, Ефименко, Хорошевский, 2011] предполагалось, что в основу будет положена некоторая обобщенная онтология бизнеса (Enterprise Ontology), которая будет коммуникативной средой между:

- разными людьми (включая пользователей и разработчиков) в различных организациях;

- людьми и компьютерными системами;

- различными реализациями компьютерных систем (включая модули самой платформы, DBMS и пр.);

а сама онтология будет служить для накопления, структурирования, представления и манипулирования бизнес-знаниями.

Успех эдинбургского проекта и создание реальных ЕКР продемонстрировали жизнеспособность такого подхода, однако массовым он не стал. Основная причина этого, на наш взгляд, состоит в трудоемкости разработки как самих баз знаний на основе онтологий, так и порталов как таковых [Загорюлько, 2009; Массель, 2010; Рыбина, 2009]. До настоящего времени этот процесс остается, скорее, искусством, чем наукой [Grosjean et al., 2011]. Формирование онтологий относится к весьма специфической форме когнитивной активности – системному структурированию и автоформализации знаний. Многочисленные исследования в этой области в основном связаны с технологиями формальных представлений и с программными средствами реализации приложений. Данная работа касается специфики разработки онтологий на начальных этапах онтологического дизайна.

Десятилетний опыт преподавания основ онтологического инжиниринга как студентам технических специальностей, так и менеджерам различной квалификации позволил автору накопить некоторый методический опыт в данной области и сформировать таксономию учебно-практических онтологий.

Если рассмотреть все множество онтологий, создаваемых в рамках одной предметной области, то можно предложить выделить 6 критериев, их характеризующих.

- Форма (описывает глубину онтологии, покрытие или объем, а также эргономические характеристики);
 - Содержание (определяется наличием или отсутствием системности, логики и уровня абстракции введенных метапонятий);
 - Фокус (определяет возможность постановки акцента на какой-либо ветви);
 - Учет времени (описывает наличие динамики);
 - Новизна (определяется оригинальностью подхода);
 - Ошибки (характеризуют синтаксис и семантику онтологии).
- Обобщенно данную новую классификацию иллюстрирует рисунок 1.



Рис. 1. Общее дерево классификаций

Эти шесть критериев позволяют предложить набор субъективных классификаций для оценки практических онтологий с учетом их зависимости от предметной области:

- А – не зависящие от предметной области,
- В – зависящие, и
- С – частично зависящие.

Далее эту классификацию можно продолжить:

А – не зависящие от предметной области

А.1.1 глубина.

А.1.1.1 глубокие

А.1.1.2 поверхностные

А.1.2 покрытие

А.1.2.1 «богатые»

А.1.2.2 скудные

А.1.3 эргономичность

А.1.3.1 гармоничные для восприятия

А.1.3.2 дисгармоничные

А.2 по содержанию

А.2.1 системность

А.2.1.1 системные

А.2.1.2 фрагментарные

А.2.2 логика

А.2.2.1 логически-оправданные

А.2.2.2 алогичные

А.2.3 генерализация

А.2.3.1 преобладание абстрактных понятий

А.2.3.2 преобладание конкретных понятий

А.2.3.3 смешанные

А.2.3.4

В – зависящие от предметной области

В.1 фокус

В.1.1 фокус на отдельных ветвях

В.1.2 сбалансированные

В.2 учет времени

В.2.1 статические

В.2.2 динамические

В.2.3

С – частично зависящие от предметной области

С.1 новизна

С.1.1 заимствованные

- C.1.2 оригинальные*
- C.2 ошибки*
 - C.2.1 корректные*
 - C.2.2 некорректные*
 - C.2.2.1 ошибки синтаксиса*
 - C.2.2.2 ошибки семантики*

При этом формула «хорошей онтологии» будет иметь вид:

- A1.1 – глубокая*
- A1.2 – «богатая»*
- A1.3 – гармоничная*
- A2.1 – системная*
- A2.2 – логичная*
- A2.3 – смешанная*
- B1 – сбалансированная*
- B2 – динамическая*
- C1 – оригинальная*
- C2 – корректная*

Только часть этих параметров поддается оценке формальными метриками (например, глубина, «богатство», баланс). Разработке инструментов формальной оценки посвящено большое количество работ [Gangemi, Catennaci, Ciaramita, & Lehman, 2006; Болотникова, Гаврилова, Горовой, 2011; и др.]. Остальные же должны оцениваться экспертами.

В практике подготовки аналитиков таким экспертом выступает преподаватель. Большинство ошибок можно охарактеризовать «дистанцией» отклонения от «хорошей» онтологии. При этом не все встречающиеся ошибки поддаются коррекции.

Так «неисправимые» относятся к классам:

- A1.1.2, A2.1, A2.2, C.*

Именно такие ошибки обусловлены особенностями когнитивных стилей разработчиков [Гаврилова, Лещева, Болотникова, Благоев, 2013]. Так, рефлексивные испытуемые строят более глубокие онтологии, таким образом, невозможно исправить характеристику A1.1. и научить импульсивного аналитика строить «глубокие» онтологии. То же относится к такой характеристике когнитивного стиля, как «Узость/широта категорий». Она влияет на ширину

онтологий: широкие категоризаторы склонны строить «широкие» онтологии.

От качества корпоративных онтологий зависит успех всего портала, потому что, сравнивая хранилище онтологий со шкафом или хранилищем предметов, можно отметить, что никто не будет пользоваться «шкафом», если полки в нем слишком узки, число створок слишком велико, а формой шкаф напоминает треугольную черепаху. Обучение же бизнес-аналитиков основам онтологического инжиниринга по-прежнему остается несбыточной мечтой директоров по информационным ресурсам и технологиям.

Таким образом, проблема разработки корпоративных онтологий, с точки зрения методологии, по-прежнему остается актуальной задачей.

СПИСОК ЛИТЕРАТУРЫ

[**Болотникова, Гаврилова, Горовой, 2011**] Болотникова Е.С., Гаврилова Т.А., Горовой В.А. Об одном методе оценки онтологий // Известия РАН, Теория и системы управления, №3. – С. 98–110.

[**Гаврилова, Лещева, Болотникова, Благов, 2013**] Гаврилова Т.А., Лещева И.А., Болотникова Е.С., Благов Е.Ю. Специфика коллективной разработки онтологий // Материалы Всероссийской конференции с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2013), Новосибирск, Россия, 2013. – С. 106–111.

[**Гурьянова, Ефименко, Хорошевский, 2011**] Гурьянова М.А., Ефименко И.В., Хорошевский В.Ф. Онтологическое моделирование экономики предприятий и отраслей современной России. Часть 2. Мировые исследования и разработки: аналитический обзор. Препринт WP7/2011/08 (ч. 2) Серия WP7. Математические методы анализа решений в экономике, бизнесе и политике. Высшая школа экономики. – Москва.

[**Загорулько, Боровикова, 2009**] Загорулько Ю.А., Боровикова О.И. Информационная модель портала научных знаний // Информационные технологии. – 2009. – № 12. – С.2–7.

[**Массель, 2010**] Массель Л.В. Применение онтологического, когнитивного и событийного моделирования для анализа развития и последствий чрезвычайных ситуаций в энергетике / Проблемы безопасности и чрезвычайных ситуаций. – №2. – 2010. – С. 34–43.

[**Рыбина, 2011**] Рыбина Г.В. Комбинированный метод приобретения знаний для построения баз знаний интегрированных экспертных систем // Приборы и системы. Управление, контроль, диагностика. 2011. №8. С. 19–41.

[Gangemi, Catennaci, Ciaramita, Lehman, 2006] Gangemi, A., Catennaci, C., Ciaramita, M., & Lehman, J. (2006). Modelling ontology evaluation and validation // In Lecture Notes in Computer Science (LNCS 4011) Ed. Sure Y. & Domingue J. The Semantic Web: Research and Applications, Springer Berlin Heidelberg. – pp. 140–154.

[Grammer, 2000] Grammer J. (2000), The Enterprise Knowledge Portal, DM Review, March 2000. – pp. 20–24.

[Grosjean, 2011] Grosjean, J., Merabti, T., Griffon, N., Dahamna, B., & Darmoni, S. (2011). Multiterminology cross-lingual model to create the European Health Terminology /Ontology Portal //In Proceedings of the 9th International Conference on Terminology and Artificial Intelligence, Paris, INALCO. – pp. 119–122.

[Uschold, King, Morales, Zorgios, 1998] Uschold M., King M., Morales S., Zorgios Y. (1998) The Enterprise Ontology. The Knowledge Engineer Review 13(1). – pp. 31–89.

УДК 004.891

МЕТОД СОЗДАНИЯ ПРОДУКЦИОННЫХ ЭКСПЕРТНЫХ СИСТЕМ НА ОСНОВЕ МОДЕЛЬНО-УПРАВЛЯЕМОГО ПОДХОДА¹

М.А. Грищенко (*makcmg@icc.ru*)

О.А. Николайчук (*nikoly@icc.ru*)

А.Ю. Юрин (*iskander@icc.ru*)

Институт динамики систем и теории управления
СО РАН, Иркутск

В работе предлагается модифицировать модельно-управляемый подход, в частности методологию MDA (Model Driven Architecture), для создания производственных экспертных систем. Произведена постановка задачи, рассмотрены основные этапы и шаблоны правил преобразования моделей.

Введение

Создание баз знаний (БЗ), а следовательно, и экспертных систем (ЭС), на этапе концептуализации знаний требует значительных усилий разработчиков. При этом исследования, направленные на повышение эффективности решения задач данного этапа охватывают ряд направлений: от совершенствования методологии (методов) разработки до создания специального инструментария, автоматизирующего процессы.

В свою очередь, при разработке прикладного программного обеспечения является перспективным использование принципов когнитивного (визуального) моделирования и подходов, основанных на преобразовании (интерпретации) информационных моделей [Гаврилова, 2000], в частности MDA (Model Driven Architecture) [Грибачев, 2004].

В данной работе предлагается модифицировать (специализировать) MDA для разработки производственных ЭС и БЗ.

¹ Работа выполнена при частичной финансовой поддержке гранта РФФИ (проект № 14-07-31299) и гранта СП-2012.2012.5 (стипендия Президента РФ).

1. Постановка задачи

Анализ существующих технологий разработки ЭС и БЗ показал, что они ориентированы в основном на пользователей, обладающих знаниями в области программной инженерии. При разработке ЭС решаются трудоемкие задачи концептуализации, формализации и моделирования предметной области на определенном языке представления знаний (ЯПЗ). При этом в большинстве случаев разработчик ЭС должен иметь навыки программирования и знать какой-либо ЯПЗ. В свою очередь, если разработку БЗ осуществляет программист, ему необходимо разобраться в огромном объеме предметных, зачастую мультидисциплинарных знаний и способов их логической обработки. Эти условия обеспечиваются в исключительно редких случаях.

В связи с этим предлагается разработать метод, основанный на модельно-управляемом подходе, обеспечивающий визуальное моделирование предметной области и автоматизированное создание продукционных БЗ и ЭС. В качестве методологической основы предлагается использовать MDA – направление в области программной инженерии, предполагающее разработку программных систем на основе трансформации и интерпретации информационных моделей, что существенно повышает качество программного средства и снижает время на его разработку.

Приведем формальную постановку задачи. Представим технологию MDA в виде:

$$MDA = \langle MOF, UML, CWM, CIM, PIM, PSM, PDM, F_{CIM-PIM}, F_{PIM-PSM} \rangle,$$

где *MOF* (Meta Object Facility) – абстрактный язык для описания информации о моделях (язык описания метамodelей); *UML* (Unified Modelling Language) – унифицированный язык моделирования; *CWM* (Common Warehouse Metamodel) – метамodel хранения данных; *CIM* (Computation Independent Model) – вычислительно-независимая модель – модель, скрывающая любые детали реализации и процессов системы, описывает только требования к системе и ее окружению; *PIM* (Platform Independent Model) – платформенно-независимая модель – модель, скрывающая детали реализации системы, зависимые от платформы, и содержащая элементы, не изменяющиеся при взаимодействии системы с любой платфор-

мой; *PSM* (Platform Specific Model) – платформенно-зависимая модель – модель системы с учетом деталей реализации и процессов, зависящих от конкретной платформы; *PDM* (Platform Description Model) – модель платформы – набор технических характеристик и описаний технологий и интерфейсов, составляющих платформу; $F_{CIM-PIM}: CIM \rightarrow PIM$, $F_{PIM-PSM}: PIM \rightarrow PSM$ – правила преобразования моделей.

Основная задача состоит в том, чтобы специфицировать (конкретизировать) методологию MDA для создания продукционных ЭС, т.е. определить MDA^{ES} :

$$MDA^{ES} = \langle MOF, UML^{ES}, CIM^{ES}, PIM^{ES}, PSM^{ES}, PDM^{ES}, F_{CIM-PIM}^{ES}, F_{PIM-PSM}^{ES} \rangle.$$

Для реализации данной задачи необходимо предложить язык для моделирования UML^{ES} и задать правила отображения моделей для обеспечения их трансформации при применении MDA для создания ЭС $F_{CIM-PIM}^{ES}, F_{PIM-PSM}^{ES}$.

Специализированная (конкретизированная) методология, реализованная в виде метода создания продукционных систем, будет апробирована в составе **программного средства**, обеспечивающего генерацию программного кода: $MDA^{ES} \rightarrow Code^{ES}$.

2. Разработка продукционных БЗ и ЭС на основе преобразования (трансформации) информационных моделей

Согласно MDA, логика работы информационной системы, включая описание основных понятий, отношений между ними и способов их обработки, представляется в виде информационных моделей, определяющих состав, структуру и поведение будущего программного продукта.

Модели и этапы создания ЭС представлены следующим образом:

- вычислительно-независимая модель (СІМ) реализована в виде онтологии: $CIM^{ES} = \langle Ont^D, Ont^{PES} \rangle$, описывающей основные понятия и отношения предметной области. При этом помимо понятий и отношений типа «является частью» и «является потомком» (иерархические связи) вводится дополнительное отношение «зависит от», обеспечивающее описание причинно-следственных зависимостей. Также описываются свойства понятий предметной области. Для

описания архитектурных элементов ЭС вводятся дополнительные понятия: «форма ввода», «форма вывода», «машина вывода»;

- платформенно-независимая модель (PIM) представлена описанием продукций, полученных путем автоматизированной трансформации СИМ и уточнения полученных результатов. При трансформации СИМ понятиям ставятся в соответствие шаблоны и элементы правил (условие и действие), причинно-следственные отношения преобразовываются в «ядро» правил. На данном этапе пользователю предлагается выбрать «начальное» правило, на основе анализа которого и всей цепочки возможного логического вывода автоматически формируется архитектура разрабатываемой ЭС, при этом отдельным экземплярам форм ввода и вывода ставятся в соответствие отдельные логические правила;

- платформенно-зависимая модель (PSM) получена в результате трансформации PIM, представлена в виде продукций на ЯПЗ JESS, их интерпретация и объяснение вывода в специальной среде-оболочке. В процессе интерпретации и объяснения вывода происходит генерация элементов пользовательского интерфейса, соответствующих PIM на языках программирования HTML, JavaScript, PHP, и их исполнение;

- модель платформы (PDM) – синтаксические правила ЯПЗ JESS и PHP, которые используются для формирования PSM на основе PIM.

Создание ЭС, помимо разработки БЗ, подразумевает разработку интерфейса взаимодействия пользователя и БЗ (или ЭС). При этом используются элементы подхода известного как ODA (Ontology Driven Architecture, поднаправление в рамках MDA) [Djuric, 2005], посвященного разработке теории и средств создания программных систем на основе трансформации онтологии. Формирование интерфейса происходит в два этапа: при создании PIM создается модель, описывающая отношения (взаимосвязи) форм ввода и вывода; при создании PSM происходит интерпретация разработанной модели на языке программирования (например, PHP).

Важным аспектом применения MDA является визуальное моделирование. При этом MDA традиционно использует UML (Unified Modeling Language) для построения моделей предметной области. Полагая, что UML не позволяет наглядно и однозначно представить причинно-следственные зависимости, предлагается при опи-

сании PIM использовать дополнительно к UML авторскую нотацию для представления продукций – RVML (Rule Visual Modeling Language) [Грищенко, 2013а].

Рассмотрим пример разработки ЭС для решения задач определения причин повреждения и разрушения материалов согласно предлагаемому методу в специализированном редакторе [Грищенко, 2013б].

Первый этап: создание CIM, включает выделение и описание понятий предметной области, а также их свойств и отношений между ними.

При решении данной задачи использована модель динамики технического состояния [Берман, 2007]. В частности, выделены следующие понятия предметной области и их свойства (рис. 1): воздействующие факторы, материал, контактная среда.

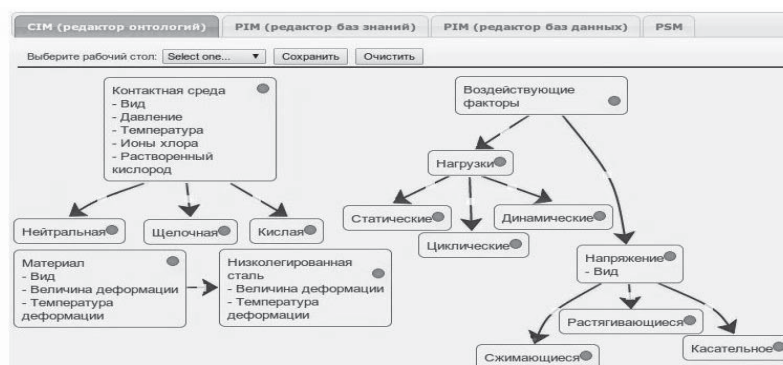


Рис. 1. Фрагмент вычислительно-независимой модели

Второй этап: создание PIM, включает создание причинно-следственных отношений в виде продукций, а также архитектуры ЭС на основе трансформации CIM. В зависимости от типа элемента понятие трансформируется в условие или факт и отображается в нотации RVML (рис. 2) для дальнейшего редактирования пользователем.

Помимо PIM для БЗ формируется модель ЭС, включающая описание основных форм пользовательского интерфейса. Данная модель формируется на основе анализа цепочки правил. В частности, в результате построения PIM были сформированы следующие правила:

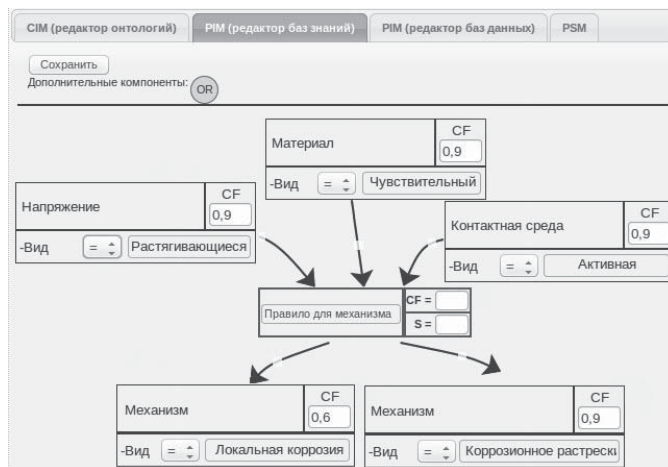


Рис. 2. Пример продукции в нотации RVML

1) **ЕСЛИ** вода под давлением 3–4 МПа и температуре около 300С⁰, ионы хлора и растворенный кислород ($K_U=1,0$) **ТО** контактная среда активная ($K_Y=0,9$).

2) **ЕСЛИ** деформация низколегированной стали в холодном состоянии величиной до 5% ($K_U=1,0$) **ТО** материал чувствительный ($K_U=0,9$) **И** остаточные напряжения.

3) **ЕСЛИ** $\sigma_{\text{раст}} \leq 0,7 \sigma_{0,2}$ ($K_U=1,0$) **ТО** значительные растягивающие напряжения ($K_Y=1,0$).

4) **ЕСЛИ** значительные растягивающие напряжения ($K_Y=0,9$) **И** чувствительный материал ($K_Y=0,9$) **И** активная контактная среда ($K_Y=0,9$) **ТО** механизм «локальная коррозия» ($K_Y=0,6$) **или** механизм «коррозионное растрескивание» ($K_Y=0,9$).

5) **ЕСЛИ** механизм – «коррозионное растрескивание» **ТО** поглощение в субмикрообъемах материала энергии остаточных напряжений и внешнего механического воздействия **И** повышение энергии кристаллической решетки до критического уровня **И** деформация кристаллической решетки **И** разрыв межатомных связей.

Правила образуют цепочку логического вывода. Для срабатывания правил не требуют дополнительной информации, используются только содержимое рабочей памяти и начальные данные, вводимые пользователем. По этой причине PIM ЭС (архитектура) включает всего 4 блока (рис. 3):

- форму ввода начальных данных для ввода информации о контактной среде, нагрузках и материале;
- машину вывода (данный элемент, являясь неотъемлемой частью ЭС, скрыт от пользователя);
- форму вывода, обеспечивающую публикацию возможных результатов логического вывода;
- форму вывода, обеспечивающую объяснение полученного решения, путем публикации списка активизированных правил и новых фактов.

Третий этап: построение PSM модели и генерация кода, включает:

- JESS/CLIPS-код БЗ (рис. 4);
- SQL-код базы данных для конкретной предметной области на основе разработанной онтологии;
- PHP и HTML код ЭС (интерфейс пользователя) (рис. 5, рис. 6).

Четвертый этап: тестирование разработанной системы, в результате которого можно либо вернуться на один из предыдущих этапов разработки и исправить неточности, либо перейти к этапу опытной эксплуатации, для опробования ЭС в рабочих условиях.

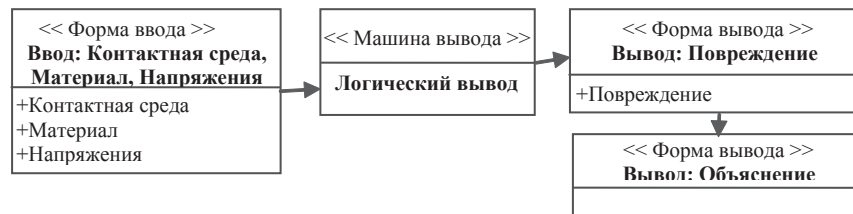


Рис. 3. PSM ЭС в нотации UML

```
(defrule pravilo-dlya-mehanizma "правило для механизма"
  (sreda ;среда
   (vid "активная")
  )
  (material ;материал
   (vid "чувствительный")
  )
  |
  (stress ;
   (vid "растягивающие") ;вид = растягивающие
  )
  => (
      assert (exist-dp "КОРРОЗИОННОЕ РАСТРЕСКИВАНИЕ")
      assert (exist-dp "ЛОКАЛЬНАЯ КОРРОЗИЯ")
    )
)
```

Рис. 4. CLIPS-код БЗ

```

<div class='dialog-form-in' id='dialog-form-newRule' title="Введите данные">
  <b>Вид контактной среды:</b>
  <textarea name="RuleName" id="RuleName" style="width: 98%" class="text ui-w
</div>
<div class='dialog-form-ln' id='dialog-form-newRule' title="Результат работы">
  <b>Коррозионное растрескивание</b> с коэффициентом уверенности = 0.9<br/>
  <b>Локальная коррозия</b> с коэффициентом уверенности = 0.6
</div>

```

Рис. 5. Пример кода экранной формы

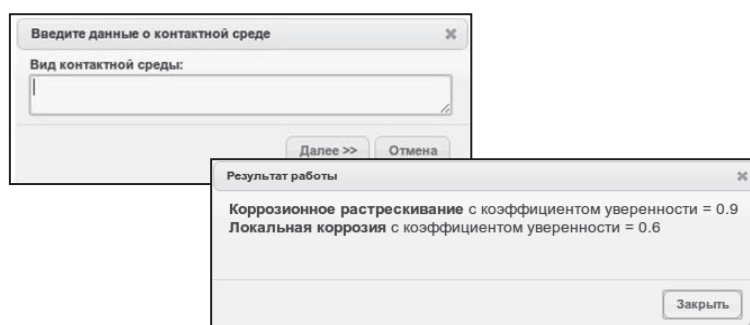


Рис. 6. Пример экранных форм ввода данных и вывода результата

Заключение

Разработан метод создания продукционных ЭС и БЗ, заключающийся в перенесении основных принципов MDA на область инженерии знаний и отличающийся от стандартного MDA использованием онтологии для представления вычислительно-независимой модели и новой графической нотации для построения моделей продукционных БЗ.

Также был проведен анализ эффективности предлагаемого метода по временному критерию. Результаты анализа показали, что эффективность разработки ЭС была повышена, в среднем на 38 % (по сравнению с классическим методом [Гаврилова, 2000]), путем:

- исключения из процесса разработки этапа реализации, за счет автоматической кодогенерации;
- уменьшения времени на этапах получения, структурирования и формализации знаний, за счет исключения из процесса разработки (или сокращения времени участия) инженера по знаниям, а также использования онтологии и когнитивной графики;

- уменьшения риска ошибок проектирования;
- исключения ошибок программирования, за счет автоматической кодогенерации.

СПИСОК ЛИТЕРАТУРЫ

[Берман, 2007] Берман А.Ф., Николайчук О.А. Пространство технических состояний уникальных механических систем // Проблемы машиностроения и надежности машин. – 2007, №1. – С.14–22.

[Гаврилова, 2000] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний и интеллектуальные системы. СПб.: Питер, 2000.

[Грибачев, 2004] К. Грибачев. Delphi и Model Driven Architecture. Разработка приложений баз данных. 2004.

[Грищенко, 2013а] Грищенко М.А., Юрин А.Ю., Павлов А.И. Программный комплекс разработки экспертных систем на основе трансформации информационных моделей предметной области / Программные продукты и системы, №3 (103), 2013. – С. 143–147.

[Грищенко, 2013б] Грищенко М.А. Разработка экспертных систем определения причин повреждения и разрушения материалов на основе трансформации информационных моделей / Материалы III Всероссийской с международным участием научно-технической конференции «Жизненный цикл конструкционных материалов (от получения до утилизации)», Иркутск: НИ ИрГТУ, 2013.

[Djuric, 2005] Dragan Djuric, Dragan Gasevic, Vladan Devedzic. Ontology Modeling and MDA. Journal of object technology, Vol. 4, No. 1, January-February 2005, pp. 109–127.

УДК 004.81, 004.89, 81.33

К СЕМАНТИЧЕСКОЙ ГЕОИНТЕРОПЕРАБЕЛЬНОСТИ В ГЕОПРОСТРАНСТВЕННОЙ SEMANTIC WEB¹

С.К. Дулин (*s.dulin@ccas.ru*)

Вычислительный центр им. А. А. Дородницына РАН, Москва

Н.Г. Дулина (*ngdulina@mail.ru*)

Вычислительный центр им. А. А. Дородницына РАН, Москва

О.С. Кожунова (*kozhunovka@mail.ru*)

Институт проблем информатики РАН, Москва

Проблема семантической геоинтероперабельности заключается в обеспечении согласованного взаимодействия специалистов для решения задач, требующих совместного использования георесурсов, при условии адекватного понимания ими семантики, заложенной в пространственные онтологии и/или геотезаурусы и классификаторы.

Работа посвящена обсуждению понятия семантической интероперабельности и аспектам ее реализации.

Введение

Возможность совместного использования геопространственных данных была одним из основных требований, начиная с разработки первой ГИС. Существующие геоданные были получены независимо друг от друга разными организациями с помощью различных систем и используются во множестве приложений также независимо друг от друга.

Несмотря на то, что прошлые два десятилетия были очень продуктивными с точки зрения развития геоинтероперабельности, облегчающей совместное использование географических данных, геоинтероперабельность пока еще не стала общезначимой. Хотя стандарты, разработанные ISO/TC 211 и Open Geospatial Consortium (OGC) Inc., обеспечили основу для ее развития и в международных организациях (Canadian Geospatial Data Infrastructure (CGDI), National Spatial Data Infrastructure (NSDI), Infrastructure for Spatial Information in Europe (INSPIRE), Global Spatial Data Infrastructure (GSDI) возникли пространственные инфраструктуры данных, гео-

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 14-07-00040).

интероперабельность все еще находится на начальной стадии реализации.

С другой стороны, за эти два десятилетия произошло внушительное развитие Интернета и Web. Организация Web началась с очень простой публикации соединенных между собой Web-страниц. Теперь Web составлена из порталов, сервисов, данных, документов, видео, музыки, постоянно пополняясь коллекциями данных различных видов типа wikis, карты Google, OpenStreetMap и т.д. В настоящее время Web совершенствуется до Semantic Web (или Web 3.0), эволюционируя от Web документов до Web данных, превращаясь в международную открытую базу данных. Все это внушает определенный оптимизм в отношении реализации геоинтероперабельности.

1. Геоинтероперабельность

Международная организация по стандартизации ISO 19119 предлагает следующее определение (http://www.wmo.int/pages/prog/www/TEM/ET-WISC-I/ISO_191xx.doc): «Интероперабельность представляет собой способность соединяться, выполнять программы или передавать данные среди различных функциональных модулей способом, который не требует, чтобы пользователь имел знания о характеристиках этих модулей».

Для географической области применимо следующее описание термина «географическая интероперабельность» (http://www.wmo.int/pages/prog/www/TEM/ET-WISC-I/ISO_191xx.doc): «Географическая интероперабельность – это способность информационных систем к 1) свободному обмену всех видов пространственной информации о Земле и об объектах и явлениях на выше и ниже поверхности Земли; и 2) совместному сетевому использованию программного обеспечения для управления такой информацией».

Следует отметить, что это определение не предполагает, что каждый компонент использует один и тот же формат данных (совпадение форматов соответствует обычно неправильному восприятию многими людьми интероперабельности), а скорее, провозглашает способность понимать формат(ы) друг друга.

Семантическая интероперабельность соответствует сотрудничеству на приближенном к человеческому уровню, и имеет дело

с потоком работ, установлением партнерства и сотрудничества, и потребностью устранить существующие (национальные, местные, организационные) барьеры. Другими словами, семантическая интероперабельность становится способностью коллекции системных компонентов к (а) распространению информации и (б) обработке ее согласно общедоступной операционной семантике, чтобы достигнуть определенной цели в данном контексте.

В контексте географической информации (ГИ) интероперабельность непосредственно связана с ГИС. ГИС представляет собой основную среду для реализации геоинтероперабельности, потому что в ней в центре внимания пространственные данные, которые более сложны, чем обычные текстовые данные, сохраненные в реляционных базах данных.

Исследования последних лет указывают на необходимость создания моделей интероперабельности, которые могут гарантировать, что интероперабельность устанавливается между системами в соответствии с различными целями и контекстами [Kingston, 2005].

Существует несколько подходов к формированию ГИС посредством моделей интероперабельности [Turnitsa, 2006]. Каждый подход имеет преимущества и недостатки относительно достижения интероперабельности в определенном контексте. Основные преимущества моделей интероперабельности – это возможность (а) определения общего словаря, который обеспечивает единообразие семантики и возможность анализа; (б) альтернативы предложений относительно структуры решений и, наконец, (с) оценки новых идей и добавления различных опций.

Примерами моделей интероперабельности, которые были успешно применены вне специфики ГИС области, являются C4ISR Architectures Working Group's Levels of IT Systems Interoperability (LISI) model (LISI 1998), the

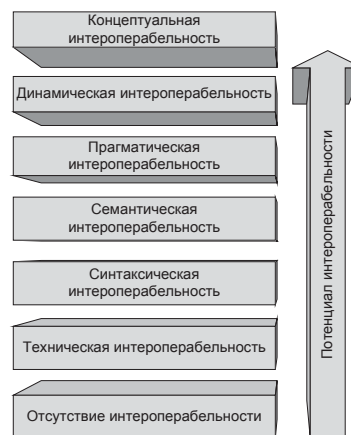


Рис. 1 Уровни интероперабельности

Enterprise Interoperability Maturity Model (EIMM) (Athena 2005), the Organisational Interoperability Maturity Model (OIMM) (Clark and Jones 1999) и the Organisational Interoperability Agility Model (OIAM) (Kingston et al. 2005). Две модели – the Levels of Conceptual Interoperability Model (LCIM) и the Intermodel5 использовались в ГИС области. Все эти модели в основном используются на самых высоких уровнях организационной интероперабельности из традиционных семи уровней: уровень нулевой интероперабельности, технический, синтаксический, семантический, прагматический, динамический и концептуальный уровни (рис. 1) [Turnitsa, 2006].

2. Семантическая геоинтероперабельность и Semantic Web

Географические базы данных и представления пользователями реальных явлений характеризуются большой гетерогенностью, которая препятствует эффективной геоинтероперабельности. Как правило, гетерогенность декомпозируется на четыре уровня: системный, синтаксический, структурный и семантический.

Установление семантической геоинтероперабельности выходит за пределы простой возможности получить доступ к информации географических баз данных на дисплее или напечатанной на бумаге. Оно требует больше времени, заранее должен быть известен точный словарь географических баз данных, чтобы получить соответствующую информацию. Но самое существенное, что установление семантической геоинтероперабельности предполагает, что пользователи и провайдеры должны иметь релевантное понимание семантики запросов и ответов. В контексте Semantic Web такая возможность становится все более и более доступной.

Моделирование семантики [Гаврилова, 2000] глубоко внедрено в структуру геоинтероперабельности и, таким образом, обеспечивает исчерпывающее описание семантической геоинтероперабельности в целом, которая лежит в основе развития семантической пространственной инфраструктуры данных и Semantic Web геопространственной информации.

Геоинтероперабельность представляет собой основу для развития и реализации пространственных инфраструктур данных (Spatial Data Infrastructures (SDIs)) [Hjelmager, 2009]. Цель SDIs состоит в

том, чтобы координировать полезный обмен и совместное использование ГИ с использованием соответствующих сервисов.

В 2001 была выдвинута идея Semantic Web [Berners-Lee, 2001]. В ее основе лежало предложение модернизировать Web от уровня документов до уровня данных и информационного моделирования.

Semantic Web требует логических утверждений, классификации понятий, формальных моделей, правил и протоколов безопасности и доверия. Архитектура Semantic Web была подробно описана во многих публикациях. Она основана [Berners-Lee, 2006] на Uniform Resource Identifiers (URI), универсальном наборе символов (Unicode), и eXtensible Markup Language (XML).

Идея создания геопространственной Semantic Web впервые была представлена в 2002 году [Egenhofer, 2009]. Она должна расширить понятие Semantic Web, улучшив семантическую функциональную совместимость ГИ в Web.

За это время усилия по стандартизации ISO/TC 211 и OGC и развитие геоинформатики обеспечили в большой степени основу для создания геопространственной Semantic Web. Международные стандарты ISO/TC 211 определили онтологию геопространственных понятий, которые являются независимыми от приложений. Эта онтология – основа описания ГИ, которая включает понятия для описания геометрии, топологии, временной информации, пространственных систем справочной информации, особенностей, характеристик, поведения, отношений, качества, метаданных и сервисов.

3. Первые шаги по реализации геоинтероперабельности

Попытки реализации некоторых аспектов геоинтероперабельности были предприняты авторами при формировании геоинформационного портала ж/д отрасли. Основное внимание при этом было уделено формированию и обработке геоописаний, учитывающих семантику, заложенную заинтересованными в их использовании лицами. Обеспечение семантической геоинтероперабельности предполагает разработку архитектуры средств согласования понимания геоданных, которая будет осуществляться на основе сравнительного анализа существующих метасхем баз геоданных с

учетом нескольких уровней взаимодействия и многофакторности взаимодействия пользователей, включая экспертов, на основе геоинформационных описаний и семантики, заложенной в пространственные онтологии и/или геотезаурусы и классификаторы. Для совместного и согласованного обмена геоданными при условии адекватного понимания пользователями семантики геоданных необходимо обеспечить хранение и сопровождение огромного объема геоинформации, включающей разнообразные тексты, документы, карты и схемы. Соответственно, одной из центральных задач здесь является извлечение релевантной геоинформации из этих геоданных с участием блока рассуждений на основе онтологий геоописаний. На рис. 2 показана схема из блоков, участвующих в интеллектуальном анализе геоданных.

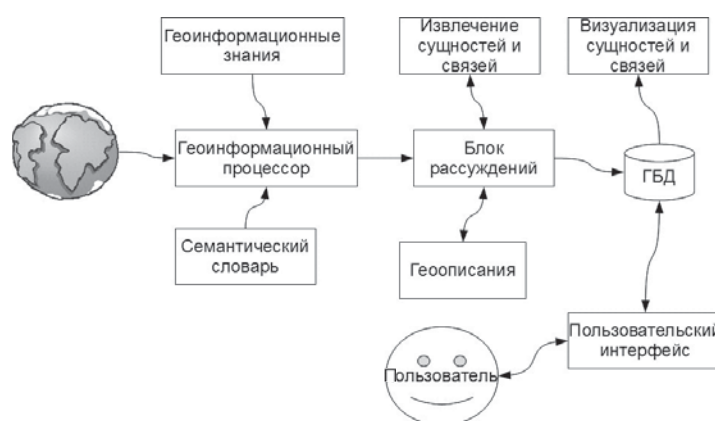


Рис. 2. Извлечение геоинформации на основе анализа геоописаний

Пространственная информационная системная среда формирования геоописаний представлена объединением ГИС с внешними программами и специальными прикладными программами, обеспечивающими выборку необходимых геоданных. Особую важность здесь представляет инструмент взаимодействия с репозиторием геоданных, который позволяет адекватно пополнять существующие геоописания за счет соотнесения с другим геоописанием того же объекта, представленным другим пользователем.

Различия имеющихся у разных пользователей геоописаний делают актуальной задачу трансфера геоописаний. Под задачей тран-

сфера геоописаний будем понимать процесс анализа и соотнесения двух или более геоописаний на основе сопоставления релевантной информации. Этот процесс является крайне необходимым при пополнении репозитория геоданных [Baird, 2005], так как позволяет получить наиболее подробное и разностороннее представление об описываемых в системе геообъектах.

Для трансфера геоописаний использовалась аналитическая машина языка функционального программирования. Следует отметить, что аналитическая машина, в основе которой лежит функциональный подход, ориентирована, в первую очередь, на анализ нормализованных геотекстов. Под нормализованным геотекстом понимается такой геотекст, который был обработан синтаксическим и/или морфологическим анализатором.

Подход к описанию правил грамматики позволил применить парадигму функционального программирования при технической (программной) реализации аналитической машины. Другими словами, имея грамматику геоязыка (заданную в виде функций) и среду функционального программирования (например, Erlang или Haskell), мы получили аналитическую машину, позволяющую проводить анализ геотекстов согласно заданной грамматике.

На представленной схеме (рис. 3) видно, что трансфер геоописаний проходит почти те же стадии обработки, что и межъязыковой трансфер. Как и любой язык, геоязык имеет синтаксис и морфологию, соответственно в системе трансфера геоописаний наличие морфологического и синтаксического анализатора является обязательным условием. После обработки указанными анализаторами получается нормализованный геотекст. Такой геотекст подается на вход системе функционального программирования, которая, используя функциональные грамматические правила, представляет геоописание в функциональном виде. Дальнейшая обработка полученного геоописания сводится либо к синтезу геоописания на другом геоязыке, либо к пополнению и обновлению данных в репозитории геоданных.

Таким образом, в результате работы системы трансфера геоописаний пользователь ГИС получает возможность соотнесения своих геоописаний с геоописаниями других пользователей, обеспечивая возможность пополнения и обновления системного репозитория геоданных.

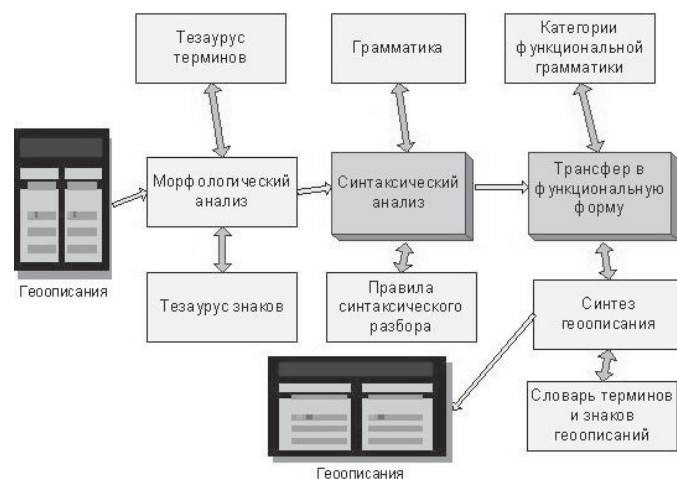


Рис. 3. Функциональная схема трансфера геоописаний

Описанный выше подход к формированию геоописаний соответствует работе между синтаксическим и семантическим уровнями (рис. 1). Проблема достижения уровня семантической интероперабельности в плане формирования согласованного набора геоописаний усложняется тем, что в большинстве реализаций ГИС до настоящего времени пространственные данные сохранялись и обрабатывались в персональных или ведомственных базах геоданных, которые ограничивали совместное использование и редактирование данных.

Заключение

Возрастающие требования к совместной обработке пространственных данных для различных приложений выявили острую потребность в масштабируемости ГИС, поддерживаемой геоинтероперабельностью, и создании пространства геоописаний, интегрированного в прикладные ГИС. Актуализация пространства геоописаний, поддерживаемого инструментальной средой ГИС, в современных условиях предполагает разработку и сопровождение централизованных репозиторий, основанных на интегрированных моделях геоданных, которые представляют различные аспекты информации жизненного цикла объектов, типа мониторинга, функционирования, восстановления и оценки стоимости.

Для достижения семантической геоинтероперабельности необходима разработка как системы согласования геоописаний, так и совокупности соответствующих сервисов для обеспечения эффективного процесса двунаправленной связи, где пользователь и поставщик ГИ взаимодействуют через запросы и ответы, понимая друг друга благодаря их знаниям и процессам рассуждения. Введение парадигмы Semantic Web открывает принципиально новые перспективы для реализации семантической геоинтероперабельности, обеспечивая широкие возможности для геоинтероперабельности ГИ типа автоматической интерпретации, рассуждения и вывода.

СПИСОК ЛИТЕРАТУРЫ

[**Дулина, 2009**] Дулина Н.Г., Уманский В.И. Структуризация проблемы улучшения пространственной согласованности баз геоданных. М.: ВЦ РАН, 2009. 40 с.

[**Kingston, 2005**] Kingston G., Fewell S., Richer W. An organisational interoperability agility model. – DSTO Fern Hill, Department of Defence, Canberra ACT 2600, Australia, 2005. [http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA463924 & Location=U2&doc=GetTRDoc.pdf](http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA463924&Location=U2&doc=GetTRDoc.pdf).

[**Turnitsa, 2006**] Turnitsa, C., and Tolk, A. Battle management language: a triangle with five sides. Proceedings of the Simulation Interoperability Standards Organization (SISO) Spring Simulation Interoperability Workshop (SIW), Huntsville, AL, 2–7 April 2006. Paper 06S-SIW-016.

[**Гаврилова, 2000**] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем // СПб.: Питер. 2000.

[**Hjelmager, 2009**] Hjelmager J., Moellering H., Cooper A., Delgado T., Rajabifard A., Rapant P., Danko D., Huet M., Laurent D, Aalders H., Iwaniak A, Abad P., Düren U., Martynenko A. An initial formal model for spatial data infrastructures, Int. J. Geogr. Inf. Sci. 22, 1295–1309, 2008.

[**Berners-Lee, 2001**] Berners-Lee T., Hendler J., Lassila O. The Semantic Web, Sci. Am. May, 34–43, 2001.

[**Berners-Lee, 2006**] Berners-Lee T., Shadbolt N., Hall W. The Semantic Web revisited // IEEE Intelligent Systems, Vol. 21. No. 3. P. 96–101, 2006.

[**Egenhofer, 2009**] Egenhofer M.J. Toward the semantic geospatial Web, GIS '02, Proc. 10th ACM Int. Symp. Adv. Geogr. Inf. Syst. ACM, New York, pp. 1–4, 2002.

[**Baird, 2005**] Baird M. P., Frome R. J. Large-scale repository design // Cell Preservation Technol., 2005. Vol. 3. No. 4. P. 256–266.

УДК 004.5

ФОРМАЛЬНАЯ МОДЕЛЬ ЛУЧШИХ ПРАКТИК ПОСТРОЕНИЯ ДОРОЖНЫХ КАРТ

И.В. Ефименко (*iefimenko@hse.ru*)
НИУ ВШЭ, Москва

В.Ф. Хорошевский (*khor@ccas.ru*)
Вычислительный центр РАН, НИУ ВШЭ, Москва

В работе обсуждаются вопросы создания формальной модели предметной области дорожного картирования на основе анализа лучших практик в данной области. Предполагается, что создание такой модели позволит обобщить опыт построения дорожных карт и обеспечит основу для разработки средств автоматизации формирования и использования дорожных карт.

Введение

Дорожное картирование, в частности, построение технологических дорожных карт (ТДК), является одним из важнейших направлений в практике Форсайт-исследований во всем мире, позволяющих учитывать специфику целей и задач прогнозирования в каждом конкретном случае, а также способствующих вовлечению в процесс достижения поставленных целей широкого круга заинтересованных лиц [Соколов, Карасев, 2009]. К настоящему моменту в России и в других странах накоплен существенный опыт создания ТДК, сформирована система понятий данной области, выработаны лучшие практики. Таким образом, может быть поставлена задача разработки формальной модели предметной области дорожного картирования. Создание такой модели, во-первых, позволит обобщить наиболее ценный опыт построения дорожных карт (ДК), во-вторых, обеспечит основу для разработки средств автоматизации формирования и использования ДК.

Ключевыми компонентами модели, представленной в рамках настоящей работы, являются:

- Система онтологий, специфицирующих ключевые информационные объекты сферы дорожного картирования – статическая составляющая дорожного картирования.

- Процессные модели лучших практик построения ДК и зон их применения с использованием упрощенных (без обеспечения высокого уровня детализации) схем алгоритмов и формализма типа «рыбий глаз» – динамическая составляющая модели;

- Гибридные (лингвистические и статистические) модели, управляющие процессами мониторинга в открытых источниках сведений об элементах ДК [Ефименко, 2012; Ефименко, 2013].

Настоящая статья посвящена описанию результатов в рамках разработки и реализации первой составляющей (система онтологий).

1. Онтология дорожного картирования: базовые компоненты

1.1. Базовая форма ДК

Как исследователи, так и практики отмечают гибкость методологии дорожного картирования, многообразие форм ДК и большое число подходов к их созданию ([Goenago-Larranaga et al., 2010; Industry Canada, 2005] и др.). Однако анализ источников и реальных проектов по построению дорожных карт ([Jurgens-Kowal, 2011; Australia, 2010; Energy, 2010; Катализаторы, 2010] и др.) позволяет выделить типовые процессы в рамках дорожного картирования и классы информационных объектов, наиболее часто выступающих в качестве элементов ДК, которые и должны быть включены в формальную модель дорожного картирования.

В рамках предлагаемой модели в качестве базовой формы используется многослойная ДК, включающая следующие слои:

- Научно-технологическое развитие (R&D);
- Технологии (Technologies);
- Продукты (Products);
- Рынки (Markets);
- Глобальный контекст развития (Global context). Подтипы элементов: Цель (Target), Глобальный тренд (Global Trend), Барьер (Barrier), Фактор риска (Risk), Окно возможностей (Opportunity).

В модели учтены ключевые характеристики (атрибуты) указанных концептов и связи между ними. Другие типы концептов, часто используемые в практике дорожного картирования, и другие способы представления дорожных карт соотносятся в модели

с множеством базовых информационных объектов через аналогии или другие типы отношений («разновидность», «расширение» и т.п.). На основании разработанной модели создан программный комплекс, пользователи которого могут вводить дополнительные (пользовательские) слои и атрибуты элементов ДК. Для обеспечения необходимого уровня гибкости в модели учитываются как описания элементов ДК и сведения, необходимые для хранения их расчетных значений, так и другие аспекты, в том числе различные измерения ДК в целом. Измерения ДК, заданные для конкретной карты, определяют границы процесса дорожного картирования в каждом случае с учетом целей построения ДК, предметной области, решаемых задач.

1.2. Основные измерения ДК. Характеристики ДК в целом

В качестве основных параметров ДК предусмотрены следующие:

- Формальные характеристики: Наименование (RM_Name); Дата создания (RM_Date); Автор/руководитель разработки (RM_Author);

- Структурные характеристики (архитектура ДК): Временной горизонт (RM_Horizon): краткосрочная (Short_Term, Near_Term), среднесрочная (Mid_Term) и долгосрочная перспектива (Long_Term). В модели предусмотрены также понятия прошлого (Past) и настоящего (Today). Особым информационным объектом в составе модели времени (частная онтология в рамках общей многоуровневой онтологии ДК) является «Образ будущего»; Количество и наименования слоев (RM_Layers);

- Формат (RM_Format): Многослойная ДК (RMF_Multiple_layers); Графы (RMF_Graphs); Инфографика (RMF_Pictures), в т.ч. Плоская диаграмма (RMF_Flow_charts); Одноуровневая/однослойная ДК (RMF_Single_layer) с подтипами;

- Основные измерения (содержательные характеристики): Предметная область (RM_Domain); Уровень (RM_Level), возможные варианты значений – Организация, Индустрия/сектор, Регион, Государство, Международное партнерство;

- Направленность (RM_Purpose).

Направленность определяется целью создания ТДК. В обсуждаемой модели предусмотрены следующие подклассы RM_Purpose [Вишневецкий, Карасев, 2010]:

- Планирование продуктовой линейки (RMP_Product_planning);
- Планирование спроса и предложения (RMP_Service_Scapabilities_planning);
- Стратегическое планирование (RMP_Strategic_planning);
- Долгосрочное планирование (RMP_Long_Range_planning);
- Управление знаниями/планирование интеллектуальных ресурсов (RMP_Knowledge_asset_planning);
- Программное планирование (RMP_Programme_planning);
- Процессное планирование (RMP_Process_planning);
- Интеграционное планирование (RMP_Integration_planning).

Также были учтены другие классификации, например, классификация на бизнес-ориентированные ТДК (тип «Бизнес»); промышленные ТДК (тип «Индустрия»); ТДК, ориентированные на цель (тип «Цель»); научно-технологические ДК для политики (тип «Политика»).

Ключевым абстрактным классом объектов является «Элемент ДК (RM_Element)». Отдельные характеристики, независимые от слоя, например, «Оценка влияния элемента на развитие предметной области ДК (RME_Impact)», «Оценка вероятности появления элемента во временном горизонте ДК (RME_Plausibility)», вынесены на уровень абстрактного элемента ДК, другие предусмотрены в рамках конкретных слоев.

В общей сложности объем сети, специфицирующей ключевые концепты практик дорожного картирования и связей между ними, составил более 500 узлов.

1.3. Характеристики классов – элементов основных слоев ДК

Ниже, в силу ограничений на объем статьи, в качестве примера представлены атрибуты класса «Научное направление» (элемент слоя «Научно-технологическое развитие»).

Базовыми атрибутами являются следующие:

- Наименование (RD_Name);
- Год появления (RD_start_Year);
- Год завершения исследований (RD_fin_Year);
- Оценочная стоимость (RD_Price);
- Необходимые ресурсы (RD_Resources);

- Уровень развития в России (RD_Rus): превышает мировой уровень, соответствует мировому уровню, незначительно / значительно отстает;

- Уровень развития за рубежом (RD_Int);

- Требования к результату (RD_Results): перечень возможных технологий, формируемых на основе направления.

Дополнительные атрибуты элемента «Научное направление» (как и других элементов ДК) предусмотрены в модели с целью обеспечения возможности детализированного описания элементов, унификации (по крайней мере, частичной) пользовательских атрибутов, а также для управления анализом документов при мониторинге открытых источников с целью извлечения сведений для предварительного заполнения элементов и связей ДК. В результате такого анализа формируется приложение к анкете эксперта, участвующего в создании ДК. В состав дополнительных атрибутов класса «Научное направление» входят следующие:

- Предметная область исследований (RD_Domain);

- Тип/Направленность (RD_Type): прикладные vs. фундаментальные исследования. Используется, например, в составе эвристик для определения значения атрибута «Уровень развития в России» при выполнении автоматического анализа (учитываются сведения о необходимости выполнения поисковой НИР vs. НИОКР и т.п.);

- Предмет исследования (RD_Subject);

- Объект исследования (RD_Object);

- Физическая основа (RD_Base): используемые вещества, материалы;

- Методы (RD_Methods);

- Оборудование (RD_Equipment);

- Этап (RD_Stage), связан с атрибутом RD_Type;

- Центры превосходства (RD_Center);

- Ключевые научные труды (RD_Papers);

- Мероприятия (RD_Events);

- Сообщества (RD_Community);

- Источники финансирования (RD_Funding);

- Объемы финансирования (RD_Fund_Vol);

- Программы/проекты (RD_Programs);

- Ведущие исследователи (RD_Researchers);

- Изобретения/патенты (RD_Invent);

- Родственные направления (RD_Rel): смежные направления исследований, направления-предшественники.

2. Онтология дорожного картирования: расширения

2.1. Компоненты модели за рамками описания основных слоев ДК

В дополнение к описанию слоев ДК и характеристик их элементов разработанная модель включает следующие классы концептов:

- Связи между элементами ДК, в частности: (1) *ограничивает (restricts)* – элемент X не может появиться раньше элемента Y (атрибут: *Оценка вероятности появления связи во временном горизонте ДК, Restriction_plausibility*); (2) *влияет (has_impact)*, атрибуты: *Сила влияния* (значения от -1 до 1), *Оценка вероятности появления связи во временном горизонте ДК*; (3) *является альтернативой (alternates)*, может использоваться при расчете основной и альтернативных технологических траекторий [Ена, Нагаев, 2013]; (4) *подкласс (is_a)*;

- Нетехнологические аспекты влияния, могут относиться к отдельному элементу, слою и ДК в целом и включают следующие подклассы: Политический аспект (Policy); Экономический аспект (Economics); Социальный аспект (Society); Экологический аспект (Environment);

- «Ценности» (Values);

- Дополнительные типы элементов / слоев ДК (наиболее частотные концепты, значимые понятия предметной области).

По результатам автоматического анализа открытых источников весовые характеристики, соответствующие различным аспектам влияния, приписываются элементам «Глобального контекста развития» (так, например, барьер «Длинные сроки строительства нефтеперерабатывающих заводов и высокие финансовые затраты» рассматривается как экономический контекст, а окно возможностей «Изменение ставок экспортных пошлин на светлые и темные нефтепродукты» характеризуется, среди прочего, политической составляющей).

Относительно дополнительных типов за рамками базовых элементов и слоев можно отметить следующее. В силу гибкости

методологии дорожного картирования перечень концептов, используемых при построении ДК, является, фактически, открытым множеством. Однако анализ источников и проектов показывает, что возможно определить совокупность концептов, наиболее часто используемых при создании ДК. Множество этих понятий, которые, таким образом, можно рассматривать как типовые элементы ДК, одновременно образует терминологическое поле дорожного картирования. В целях установления формальных взаимосвязей (отображений) базовых и дополнительных элементов и слоев ДК в модели предусмотрено несколько различных политик. Включение таких концептов в модель позволяет создать инструментарий, обладающий необходимым уровнем выразительной мощности, и одновременно обеспечить унификацию описания. Примерами таких концептов являются: *Рыночный, технологический и социальный тренды, Драйвер, Ресурс, Слабый сигнал, Форс-мажорный фактор (джокер), Угроза, Лакуна, Разрушающая технология (Disruptive technology)* и т.п.

2.2. Дополнительные форматы представления ДК.

Система онтологий

В рамках предложенной модели, как указывалось выше, в качестве базового выбрано представление в формате многослойной ДК с заданным набором слоев, ориентированным на тип «Планирование продуктовой линейки» (RMP_Product_planning). Однако система онтологических моделей обеспечивает возможность описания не только выбранной базовой формы представления ДК, но и дополнительных форматов. При этом устанавливается соответствие между слоями базового формата и слоями других форматов ДК. Учтены следующие форматы: «*Планирование спроса и предложения*», «*Стратегическое планирование*», «*Управление знаниями/планирование интеллектуальных ресурсов*», «*Интеграционное планирование*» [Вишневский, Карасев, 2009]. Пример для формата «*Управление знаниями/планирование интеллектуальных ресурсов*» представлен ниже (рис. 1, табл. 1).

В дополнение к онтологии дорожных карт как таковых в системе онтологий представлены следующие модели: «Онтология экспертов и экспертных процедур», «Онтология источников информации», вспомогательные модели – «Онтология лингвистических

шкал», онтологии предметных областей «Стандарты и регулирование», «Производство и внедрение» и др. Вспомогательные модели предназначены, прежде всего, для автоматизации выявления в материалах открытых источников сведений об элементах ДК и других информационных объектов научно-технического прогнозирования.

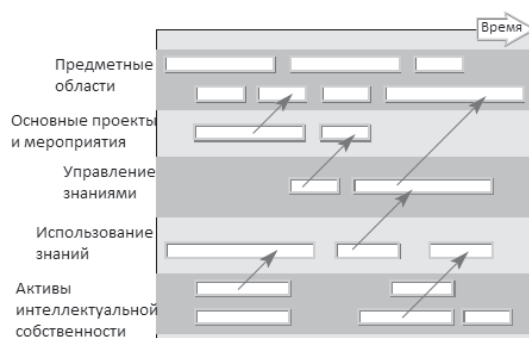


Рис. 1. Формат «Управление знаниями»

Таблица 1

Сопоставление базовой и дополнительных форм

Управление знаниями	Базовая форма
Предметные области (Business objectives)	Рынки
Проекты и мероприятия (Leading projects & actions)	Продукты
Управление знаниями (Knowledge management enablers)	Технологии; Драйверы
Использование знаний (Knowledge related processes)	Технологии
Активы (Knowledge assets)	Научно-технологическое развитие

Фрагмент итоговой системы онтологий в редакторе Protégé представлен на рис. 2.

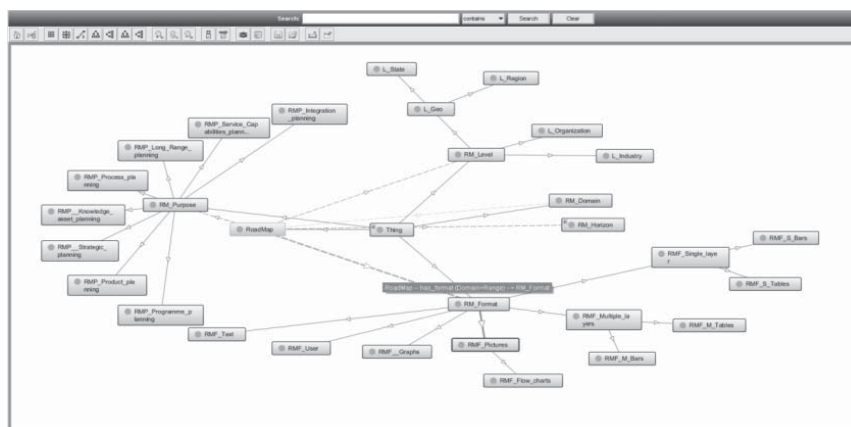


Рис. 2. Фрагмент онтологии дорожного картирования

Заключение

В статье представлены результаты, полученные при разработке системы онтологий технологических дорожных карт на основе анализа лучших российских и зарубежных практик. Разработанная система онтологий может быть использована в качестве основы информационно-аналитических систем в области научно-технологического прогнозирования. Ключевые направления дальнейших работ – создание модулей автоматической обработки неструктурированной информации, ориентированных на извлечение знаний обо всех ключевых информационных объектах дорожного картирования, а также подготовка альбома методологических схем для разработчиков дорожных карт.

СПИСОК ЛИТЕРАТУРЫ

- [Соколов, Карасев, 2009] Соколов А.В., Карасев О.И. Форсайт и технологические дорожные карты для nanoиндустрии. Российские нанотехнологии, Т. 4, № 3–4, 2009.
- [Industry Canada, 2005] Industry Canada. Technology Roadmapping: A Guide for Government Employees. Ottawa: Industry Canada, 2005.
- [Goenago-Larranaga et al., 2010] Goenago-Larranaga, JM and Phaal, R. Roadmapping in industrial companies: Experience. DYNA-BILBAO, 85. pp. 331–340. ISSN 0012-7361, 2010.

[Jurgens-Kowal, 2011] Jurgens-Kowal, Teresa. Roadmapping 101. Directing Technologies to Win the Future. Global NP Solutions, LLC, 2011.

[Australia, 2010] Automotive Australia, 2010. The Automotive Australia 2020 (AA2020) roadmap project, <http://www.autocrc.com/about/2020>.

[Energy, 2010] Energy Technology Roadmap: a Guide to development and Implementation. OECD/IEA, 2010.

[Катализаторы, 2010] Дорожная карта «Использование нанотехнологий в каталитических процессах нефтепереработки», 2010, <http://www.hse.ru/org/hse/tecnpl/dk>.

[Вишневский, Карасев, 2010] Вишневский К.О., Карасев О.И. Прогнозирование развития новых материалов с использованием методов Форсайта //Форсайт, Т. 4, № 2, 2010.

[Ена, Нагаев, 2013] Ена О.В., Нагаев К.В. Автоматизация процессов создания и применения технологических дорожных карт. Расчет интегральных показателей применимости. Бизнес-информатика. 2013. № 3 (25). С. 20–26.

[Ефименко, 2013] Ефименко И.В. Гибридный подход к выявлению комплексных объектов в области научно-технического прогнозирования: принцип «черного ящика». В кн.: Открытые семантические технологии и системы, OSTIS-2013 (г. Минск, Беларусь), БГТУ, 2013.

[Ефименко, 2012] Ефименко И. В. Модели использования и интерпретации оценочной информации в прогнозировании: время, состояние, вероятность. В кн.: Тринадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2012 (г. Белгород, Россия), Том 1. Российская ассоциация искусственного интеллекта, 2012. С. 244–251.

УДК 004.8

ПОДДЕРЖКА УПРАВЛЕНИЯ БАЗАМИ ЗНАНИЙ ПРИ АВТОМАТИЗАЦИИ ИНТЕЛЛЕКТУАЛЬНОЙ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ¹

А.С. Клещев (*kleschev@iacp.dvo.ru*),

Е.А. Шалфеева (*shalf@iacp.dvo.ru*)

Институт автоматизации и процессов управления ДВО РАН,
Владивосток

Статья посвящена организации процесса управления качеством баз знаний, используемых для поддержки принятия решений в повседневной интеллектуальной деятельности. Представлены схемы автоматизации процесса управления качеством баз знаний и основные задачи программных систем для управления качеством баз.

Введение

Одним из средств повышения качества профессиональной деятельности, в том числе и интеллектуальной, является ее автоматизация с помощью инструментов и технологий, таких как: СУБД, средства создания корпоративных систем, сети передачи данных, облачные вычисления, технологии разработки экспертных систем – систем, основанных на знаниях. Несмотря на значительные успехи в области теории и технологии, системы, основанные на знаниях, не получили заметного практического использования так, чтобы быть встроенными в организационную структуру учреждений и оказывать систематическую поддержку в повседневной интеллектуальной деятельности специалистов. Наиболее вероятным объяснением сложившейся ситуации является недостаточное внимание поддержке процесса управления качеством баз знаний, используемых для поддержки такой деятельности [Клещев и др., 2013].

Цель работы – показать необходимость программных систем для управления качеством баз знаний и установить основные задачи этих систем.

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 13-07-00024-а и № 12-07-00179-а) и КПФИ «Дальний Восток».

1. Базы знаний для поддержки интеллектуальной деятельности и их оценки

Под *интеллектуальной деятельностью* в работе понимается деятельность, состоящая в принятии взаимосвязанных решений на основе знаний по отношению к некоторым объектам действительности. Специфика принятия решений на основе знаний состоит в том, что алгоритмы такого принятия решений неизвестны; известны лишь алгоритмы «применения знаний» для принятия решений, при этом «качество» (правильность и точность) решений зависит от *качества знаний* [Клещев и др., 2013].

Управлением знаниями назовем деятельность, связанную с расширением, уточнением, совершенствованием знаний. Управленческие меры по *уточнению знаний* (по мере выявления в них неточностей) и *расширению знаний* (внедрению новых апробированных научных достижений) могут выполняться научными организациями, высокотехнологичными учреждениями или их отдельными специалистами.

Поддержка повседневной интеллектуальной деятельности означает принятие специалистом решения с учетом объяснения экспертной системы (ЭС). Объяснение содержит информацию о соответствии гипотез о вариантах решения задачи рассматриваемой ситуации. Специалист (например, врач), принимая решение (например, ставя диагноз) на основании данных о ситуации (например, истории болезни пациента), руководствуется своими знаниями, но получает возможность учесть формируемое экспертной системой объяснение (анализ возможных гипотез о диагнозе для этого пациента).

Каждое решение, принимаемое в ходе интеллектуальной деятельности специалистом, обычно в дальнейшем проходит процедуру верификации (которая может потребовать определенного времени). В результате этой процедуры решение признается правильным или ошибочным (речь идет о решении специалиста или решении специалиста, принятом при поддержке системы).

ЭС проводит независимый анализ входных данных на соответствие их тем или иным гипотезам о решении задачи, используя базу знаний. Если база знаний имеет высокое качество, ЭС позволяет снизить долю ошибок специалистов, связанных с неправильным применением знаний. Экспертная система всегда применяет базу знаний правильно (правильно использует правильный алгоритм

решения задачи) и может провести полный анализ любого множества гипотез. Однако результаты анализа в значительной степени зависят от качества применяемых в этом анализе знаний.

Базы знаний для ЭС, как правило, сформированы экспертами [Гаврилова, 2001; Джарратано, 2007]. Созданная экспертом база знаний может содержать дефекты – быть неполной (некоторые варианты могут быть упущены экспертом или быть ему неизвестны), неточной (может приводить к неоднозначным решениям) или даже неправильной (из-за заблуждений, предубеждений). Качество и полезность базы знаний определяются полнотой, точностью и правильностью содержащихся в ней знаний. Очевидна необходимость более объективного оценивания качества баз знаний, поскольку из литературы следует, что до сих пор основными средствами оценивания баз знаний являются средства контроля формальных свойств правильно построенной БЗ и привлечение экспертов для оценки «решений, предлагаемых системой» [Гаврилова, 2001; Соловьев и др., 1986; Тельнов, 2004].

Оценка правильности БЗ может определяться множеством задач, которые ЭС правильно решает на основе БЗ. *Оценка точности БЗ* может определяться подмножеством этого множества, для которого решения ЭС однозначны.

Если оценкой знаний специалиста считать множество задач, для которых он знает правильное решение, то ЭС можно считать полезной для некоторого специалиста, если оценка точности ее БЗ лучше оценки знаний специалиста, т.е. если множество задач, решение которых известно этому специалисту, является подмножеством множества точно решаемых экспертной системой задач. Определяемая таким способом оценка правильности и точности базы знаний зависит от множества задач с известным решением (базы прецедентов).

2. Поддержка управления качеством знаний и решений

Со временем (в процессе практики специалистов) множество задач, решение которых известно специалисту, будет расширяться. Каждое увеличивающееся со временем множество задач с известным решением (прецеденты) используется для автоматизированной проверки того, насколько база знаний удовлетворяет накопленным прецедентам. В доступных источниках приводятся при-

меры программного инструментария для обслуживания такой базы прецедентов (пополнения и модифицирования) и ее использования [Попов и др., 1996; Соловьев и др., 1986].

Зафиксированная БЗ со временем будет устаревать в том смысле, что оценка ее правильности и точности не будет изменяться, в то время как база прецедентов будет расширяться в процессе дальнейшей практики. Т.е. из-за того, что оценка специалиста улучшается, а оценки правильности и точности такой БЗ не изменяются, ЭС с такой БЗ может потерять свою полезность. Естественно ожидать, что база знаний должна со временем совершенствоваться.

Чтобы ЭС оставалась полезной для специалиста, оценка правильности и точности ее БЗ в течение всего времени эксплуатации ЭС должна оставаться лучше, чем оценка знаний специалиста.

Поэтому ЭС должна включать *систему управления базой знаний*, цель которой – обеспечивать полезность ЭС для специалиста в течение всего времени ее эксплуатации [Клещев и др., 2010].

Для достижения этой цели система управления БЗ должна выполнять следующие функции:

- постоянно накапливать базу прецедентов;
- классифицировать все прецеденты в этой базе;
- находить все возможные способы модификации БЗ для «включения» новых прецедентов в оценку правильности и точности БЗ (т.е. способы улучшения ее правильности и точности за счет охвата этих новых прецедентов);
- модифицировать БЗ одним из найденных способов.

Для того чтобы ЭС оставалась полезной для специалиста, необходимо, чтобы информация обо всех задачах, решенных этим специалистом, была доступна ЭС. Результат верификации каждой решенной задачи также должен быть доступен ЭС. Получение этой информации и является функцией накопления базы прецедентов.

Ее естественной реализацией является интеграция ЭС с электронным документооборотом интеллектуальной деятельности: информация о принятых специалистами решениях, а также о результатах верификации этих решений также должна включаться в документы в форме, допускающей ее обработку системой управления БЗ. Такой электронный документооборот позволяет получать объективные оценки качества решения задач как отдельными специалистами, так и их группами.

Каждый прецедент должен быть отнесен системой управления БЗ к одному из следующих классов (рис. 1):

- 1) ЭС предложила правильное и точное решение.
- 2) ЭС предложила правильное, но неточное решение (несколько возможных альтернатив, среди которых было и правильное решение), но входные данные задачи допускают ее точное решение.
- 3) ЭС предложила правильное, но неточное решение (несколько возможных альтернатив, среди которых было и правильное решение), но входные данные задачи допускают некоторое его уточнение (уменьшение числа альтернатив).
- 4) ЭС предложила правильное, но неточное решение (несколько возможных альтернатив, среди которых было и правильное решение), но входные данные задачи не допускают его уточнения.
- 5) ЭС предложила неправильное решение (множество альтернатив, возможно пустое, среди которых не было правильного решения).

Функция классификации прецедентов состоит в отнесении каждого прецедента к одному из указанных классов. Очевидно, что решение об отнесении прецедента к классам 2–4 не может быть принято автоматически, поэтому оно должно приниматься экспертами, входящими в «группу управления БЗ».

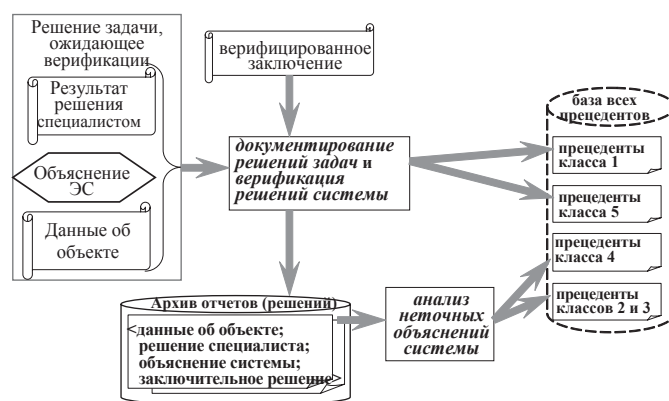


Рис. 1. Схема связи документооборота с поддержкой системы управления

На рис. 1 показано, что документируются не только результат решения задачи и все известные данные об объекте, но и объясне-

ние, полученное от ЭС. Через некоторое время становится известно и верифицированное заключение. При наличии решения задачи, сделанного специалистом, объяснения, сформированного системой, и верифицированного заключения ответственные за качество знаний лица (эксперты) могут сравнить их. В случае несовпадения этих решений следует провести анализ сгенерированного объяснения, чтобы оценить, было ли объяснение неправильным, и следует ли пару <известные данные об объекте, верифицированное заключение> использовать как прецедент для исправления базы знаний системы.

Прецеденты, отнесенные к классам 1 и 4, образуют оценку правильности и точности БЗ. Новые прецеденты, отнесенные к этим классам, могут быть включены в эту оценку без модификации БЗ, в отличие от новых прецедентов, отнесенных к классам 2, 3 и 5. Допустимой является такая модификация БЗ, которая не ухудшает ее оценку, т.е. классы, к которым отнесены входящие в нее прецеденты после допустимой модификации БЗ, не изменяются, или некоторые прецеденты из класса 4 переходят в класс 1. Новые прецеденты из классов 2 и 3 требуют уточнения БЗ, т.е. такой ее допустимой модификации, при которой прецеденты из класса 2 переходят в класс 1, а прецеденты из класса 3 переходят в классы 1 или 4; уточнение БЗ имеет целью включить эти прецеденты в оценку правильности и точности уточненной БЗ. Новые прецеденты из класса 5 требуют исправления или расширения БЗ, т.е. такой ее допустимой модификации, при которой прецеденты из класса 5 переходят в классы 1 или 4; исправление или расширение БЗ также имеет целью включить эти прецеденты в оценку правильности и точности исправленной или расширенной БЗ. В поиске всех возможных вариантов таких допустимых модификаций БЗ для всех новых прецедентов и состоит функция поиска возможностей включения новых прецедентов в оценку БЗ. Эти варианты допустимых модификаций БЗ должны вычисляться системой управления БЗ автоматически.

В результате выполнения предыдущей функции может быть получено несколько вариантов допустимых модификаций БЗ для новой группы прецедентов, либо такие варианты могут вообще отсутствовать. Поэтому функция модификации БЗ состоит в выборе

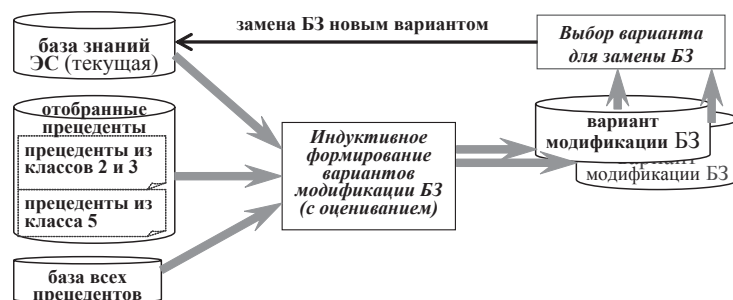


Рис. 2. Поддержка автоматического управления качеством баз знаний

одного такого варианта допустимой модификации (если они есть) и его выполнения, либо в пересмотре некоторых ранее принятых решений при модификации БЗ (если таких вариантов нет). Реализация этой функции должна осуществляться экспертами, входящими в группу управления БЗ при поддержке системы управления. Для поддержки получения формализованных новых знаний (рис. 2) требуются: удобные для эксперта средства формирования обучающей выборки из прецедентов классов 2 и 3, средства автоматического формирования очередного варианта модификации БЗ и его оценивания. Если получен вариант БЗ с оценкой не хуже оценки специалиста, он становится новой БЗ ЭС (вместо «текущей БЗ», используемой на этот момент).

Дополнительным источником совершенствования знаний, используемых при решении задач интеллектуальной деятельности, являются новые научные результаты, относящиеся к этой интеллектуальной деятельности. Естественно, что система управления БЗ должна допускать включение новых научных результатов в БЗ (без ухудшения ее оценки). Такая модификация БЗ может выполняться только экспертами, входящими в группу управления БЗ (рис. 3). Требуется автоматизированная поддержка оценивания варианта изменяемой (по результатам научных исследований) БЗ, чтобы убедиться, что оценка базы знаний улучшилась. Если оценка модифицированной БЗ становится не хуже при включении в нее новых научных знаний, то такой вариант модификации становится новой БЗ ЭС.

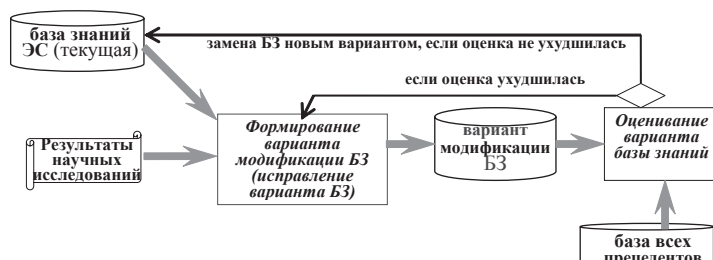


Рис. 3. Включение в базу знаний результатов научных исследований

Управленческие меры, направленные на контроль принимаемых специалистами решений и контроль используемых знаний, естественно, связаны с мерами, направленными на приведение в соответствие знаний специалистов с обновляемыми знаниями. Для поддержки обучения или повышения квалификации специалистов необходимы программные средства и компьютерные тренажеры, знания в которых полностью соответствуют самым современным формируемым базам знаний.

3. Преимущества автоматизации управления знаниями

Проектирование программных средств поддержки управления БЗ имеет преимущества как для выполняющих повседневную интеллектуальную деятельность специалистов, так и для управляющих этой деятельностью.

Преимущества эти таковы:

- получение базы знаний ЭС с оценкой качества не хуже, чем оценка качества знаний специалистов, которые пользуются этой ЭС, и достижение монотонного роста оценки качества БЗ ЭС;
- явная фиксация системы знаний с наилучшей оценкой, ее доступность для изучения (в качестве справочника), для использования на практике (в качестве стандарта) и в обучении (в качестве учебного пособия). Наличие такой системы знаний может значительно расширить возможность своевременного доведения новейших знаний до специалистов;
- объективность оценивания качества работы специалистов, если средства документооборота интегрируются с ЭС. Тогда можно оценивать качество работы отдельных специалистов или коллектива учреждения.

Заключение

Для постоянного повышения качества принимаемых решений параллельно с автоматизацией повседневной интеллектуальной деятельности должно осуществляться управление качеством баз знаний. Управление качеством баз знаний требует автоматизации и интеграции с документооборотом и ЭС. Основные задачи программных систем для управления качеством БЗ: документирование результата решения задачи с исходными данными, объяснением ЭС и верифицированным заключением, накопление и классификация новых прецедентов, поддержка сравнения решения и объяснения с верифицированным заключением, поддержка анализа объяснения, поддержка формирования вариантов допустимых модификаций БЗ для новой группы прецедентов.

СПИСОК ЛИТЕРАТУРЫ

- [Гаврилова, 2001] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001. 384 с.
- [Джарратано, 2007] Джарратано Д., Райли Г. Экспертные системы: принципы разработки и программирование. – Киев: Вильямс, 2007. 1147 с.
- [Клещев и др., 2010] Клещев А.С., Грибова В.В. Управление интеллектуальными системами. Известия РАН // Теории и системы управления. 2010. № 6. С. 122–137.
- [Клещев и др., 2013] Клещев А.С., Черняховская М.Ю., Шалфеева Е.А. Парадигма автоматизации интеллектуальной профессиональной деятельности. Часть 1. Особенности интеллектуальной профессиональной деятельности // Онтология проектирования. 2013. №3(9). С. 53–69.
- [Попов и др., 1996] Попов Э.В., Фоминых И.Б., Кисель Е.Б., Шапот М.Д. Статические и динамические экспертные системы. М.: Финансы и статистика, 1996.
- [Соловьев и др., 1986] Соловьев С.Ю., Соловьева Г.М. Методы отладки баз знаний в системе ФИАКР // Сб. Автоматизация и роботизация производства с применением микропроцессорных средств. Кишинев, 1986. С. 36–37.
- [Тельнов, 2004] Тельнов Ю.Ф. Интеллектуальные информационные системы // Московский государственный университет экономики, статистики и информатики. – М.: МЭСИ, 2004. 246 с.

УДК 681.3

ИНТЕГРАЦИЯ РЕЛЯЦИОННЫХ ДАННЫХ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО ПОДХОДА¹

А.О. Колесов (*alexander.kolesov90@gmail.com*)

А.М. Наместников (*nam@ulstu.ru*)

Ульяновский государственный технический университет, Ульяновск

В данной статье описываются принцип работы, архитектура и реализация системы интеграции реляционных данных на основе онтологического подхода. Также приведен пример использования системы на основе экспериментальных данных.

Введение

На сегодняшний день самыми популярными и распространенными источниками являются базы данных, основанные на реляционной модели. Несмотря на то, что реляционные базы данных разрабатываются с использованием общепринятых стандартов, их интеграция не является тривиальной задачей из-за возможной сложности их структуры. Так, две базы данных, содержащие записи о сущностях двух схожих предметных областей, могут быть спроектированы на разных уровнях абстракции, таблицы могут иметь разную степень нормализации, домены и ограничения целостности могут иметь различные формулировки. Помимо этого возникает проблема семантической гетерогенности, связанная с наличием разных понятий предметной области и их интерпретаций. Понятие семантической гетерогенности достаточно полно представлено в работе [Buccella et al., 2005].

В процессе интеграции реляционных источников данных на основе онтологического подхода можно выделить две основные задачи: преобразование реляционных данных в RDF формат и выделение концептуальной схемы; объединение онтологии для предоставления унифицированного доступа к данным из разных источников.

¹ Работа выполнена при финансовой поддержке РФФИ (проект №14-01-31086).

1. Отображение реляционных данных и объединение онтологий

27 сентября 2012 года консорциум W3C опубликовал документы, в которых представлены официальные рекомендации для спецификаций языков отображения DM (Direct Mapping) и R2RML (RDB to RDF Mapping Language). Согласно этим спецификациям язык DM представляет собой, как видно из названия, автоматическое прямое отображение реляционных таблиц. Структура результирующего RDF-графа при прямом отображении базы данных напрямую зависит от структуры базы (например, названия таблиц преобразуются в имена классов, а названия столбцов преобразуются в названия свойств). При этом ни структура, ни целевой словарь не могут быть изменены [Bertails et al., 2012]. Само по себе прямое отображение приемлемо только в самых простых случаях.

Реляционные базы данных на современных предприятиях и в организациях могут иметь очень сложную структуру, так как она напрямую зависит от уровня компетентности ее разработчика и от предметной области. Так, например, в базах может содержаться не одна сотня таблиц, в которых хранятся сущности реальной предметной области, в то время как другие таблицы носят лишь вспомогательный характер. Для построения отображения для таких случаев была разработана спецификация R2RML.

В отличие от прямого отображения отображение с использованием спецификации R2RML является полностью настраиваемым. Помимо этого оно позволяет работать с SQL-представлениями таблиц, посредством встраивания SQL-запросов прямо в файл с отображением [Das et al., 2012].

Языки DM и R2RML не исключают, а дополняют друг друга. Сначала пользователь с помощью DM может увидеть, что из себя будет представлять RDF-граф, и получить предварительную версию файла отображения. Затем, используя R2RML, пользователь может произвести тщательную настройку отображения для дальнейшей работы.

В работе исследователей из Цюрихского университета [Hert et al., 2011] предпринята попытка сравнения между собой существующих языков отображения реляционных данных в RDF-множества, посредством анализа индивидуальных функций и особенностей,

которые присутствуют в этих языках. В результате исследуемые языки были разделены на четыре группы: языки прямого отображения, языки общего назначения (чтение), языки общего назначения (чтение и запись), языки специального назначения. Проанализировав указанную работу, а также работу [Villazón-Terrazas et al., 2012], в которой представлены наиболее известные реализации языков DM и R2RML и результаты тестирования их работоспособности, можно выделить фреймворки XSPARQL, db2tripels и D2RQ. Данные фреймворки имеют ряд преимуществ перед всеми остальными (одновременная поддержка DM и R2RML, реализация на языке Java, открытый исходный код).

В контексте интеграции данных онтология представляет собой словарь, дающий формальное описание предметной области, которая является результатом интеграции. Таким образом, онтология выступает в качестве семантической надстройки над источниками данных и будет предоставлять унифицированный интерфейс доступа. Выделяют следующие подходы при интеграции данных:

1) Подход с использованием единой онтологии. Все схемы источников данных напрямую связаны с общей глобальной онтологией. Интеграция при использовании данного подхода является чувствительной к изменениям в информационных источниках. В зависимости от природы этих изменений в одном источнике они могут вызвать изменения в глобальной онтологии. Схема показана на рис. 1а.

2) Мультионтологический подход. Каждый источник данных отдельно описывается своей локальной онтологией. Далее устанавливаются соответствия между локальными онтологиями нескольких ресурсов. Такой подход позволит упростить процедуру внесения изменений в систему (например, добавление, удаление, модификация источника данных). Однако отсутствие единого словаря предельно усложняет процесс сравнения различных локальных онтологий. Схема показана на рис. 1б.

3) Гибридный подход. Используется комбинация двух предыдущих подходов. Строятся локальные онтологии для каждого источника данных. Затем устанавливаются соответствия, но не с другими локальными онтологиями (как в мультионтологическом подходе), а с глобальной онтологией. Схема показана на рис. 1в.

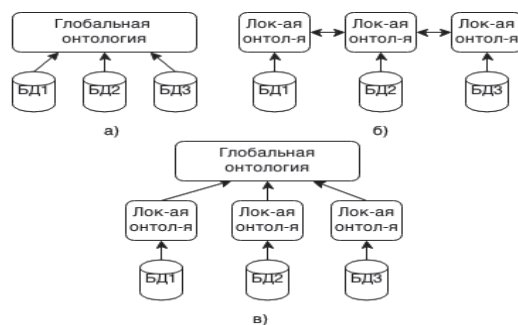


Рис. 1. Три подхода применения онтологий

В работе [Калиниченко, 2012] описывается подход доступа к реляционным данным, основанного на применении онтологии при помощи использования дескриптивной логики. В данном случае происходит преобразование конъюнктивного запроса в терминах онтологии в запрос над реляционной базой данных.

2. Реализация системы интеграции

Разработанная система предоставляет единый интерфейс для работы с реляционными данными из нескольких источников посредством использования единого словаря или, другими словами, онтологии. Структурная схема системы интеграции показана на рис. 2.

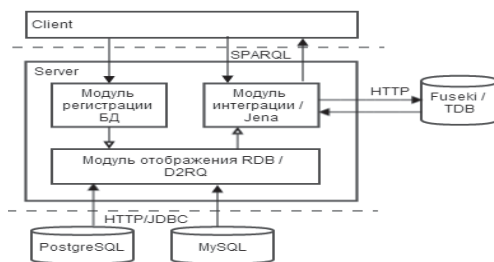


Рис. 2. Структурная схема системы интеграции

Система реализована с помощью объектно-ориентированного языка программирования Java на платформе Java EE, т.к. она имеет развитые методики и подходы в быстрой разработке Web приложений. Использование виртуальной машины Java как среды выполнения позволит встраивать в систему блоки кода, написанные

на таких многообещающих языках, как Groovy, Scala или Clojure. Что в свою очередь значительно увеличивает количество вспомогательных библиотек, которые можно использовать в системе. Так, например, появляется возможность использования таких реализаций языка R2RML как Morph (реализован на языке Scala) или Clj-R2RML (реализован на языке Clojure).

Процесс интеграции данных при использовании данной системы состоит из следующих этапов:

1) Регистрация сведений о базах данных, таких как имя пользователя, пароль и адрес сервера, на котором расположена БД. Данный шаг реализован с помощью соответствующего модуля.

2) Извлечение RDF и RDFS данных из каждой базы. Этот этап осуществляется с помощью модуля отображения. Данный модуль реализован на основе API библиотеки D2RQ, которая поддерживает такие языки отображения реляционных данных в RDF тройки, как DM и R2RML. Выбор сделан в пользу D2RQ, так как эта библиотека имеет встроенные методы для работы с фреймворком Jena.

3) Предварительная обработка и сохранение извлеченных данных в хранилище RDF троек. Выполнение данного этапа осуществляется с помощью модуля интеграции, в основе которого лежит фреймворк Jena. В качестве хранилища выступает SPARQL сервер Fuseki. Одним из преимуществ этого сервера является наличие готового графического интерфейса. Данная особенность является очень полезной при проведении экспериментов, особенно на ранней стадии разработки основной системы.

4) Объединение словарей (онтологий) при помощи добавления связей типа `rdfs:SubClassOf` и `owl:equivalentClass`.

5) Построение запроса с учетом структуры объединенного словаря и получение необходимой информации из RDF хранилища с использованием языка запросов и протокола SPARQL [Prud'hommeaux et al., 2008]. Язык SPARQL обладает гибкостью и выразительностью, что позволяет строить запросы любой сложности.

3. Вычислительные эксперименты

Для проверки работоспособности системы проведены вычислительные эксперименты. В качестве исходных данных для экспериментов были разработаны две базы данных. Разработка велась с помощью двух наиболее популярных реляционных СУБД, Postgres

и MySQL, соответственно. Для проведения вычислительных экспериментов серверы баз данных были установлены на разные рабочие станции, соединенные между собой локальной сетью. В первой базе данных хранятся записи о научных статьях и связанные с ними данные, такие как авторы, разделы, ключевые слова, список литературы и т.д. База содержит 1500 записей об авторах, 2000 записей о статьях, 230 разделов, 50 ключевых слов, 500 записей о списках литературы. Фрагмент схемы базы показан на рис. 3.

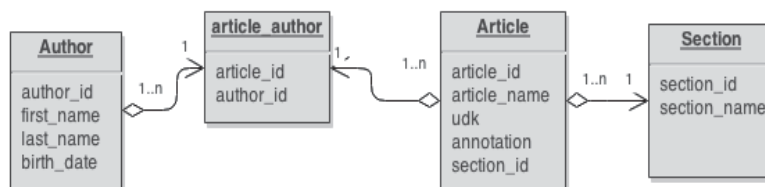


Рис. 3. Фрагмент схемы базы данных научных статей

Во второй базе данных содержится информация о сотрудниках университета. В базе хранятся 500 записей о сотрудниках и 18 записей о структурных подразделениях, в которых они числятся. Среди всех сотрудников только 1/5 часть является авторами научных статей. Фрагмент схемы базы показан на рис. 4.

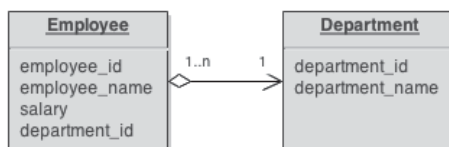


Рис. 4. Фрагмент схемы базы данных сотрудников университета

В результате работы модуля отображения из каждой базы данных были сформированы соответствующие RDF тройки, которые можно представить в виде графов. В рамках эксперимента предположим, что все авторы научных статей являются сотрудниками университета. Но так как в университете есть подразделения, не связанные с научной работой, то, следовательно, не все сотрудники университета могут быть авторами.

Таким образом, мы можем установить связь между графами посредством добавления свойства `rdfs:SubClassOf`. После установки

связи общее число RDF троек составило 36240. Результат объединения графов в упрощенном виде показан на рис. 5.

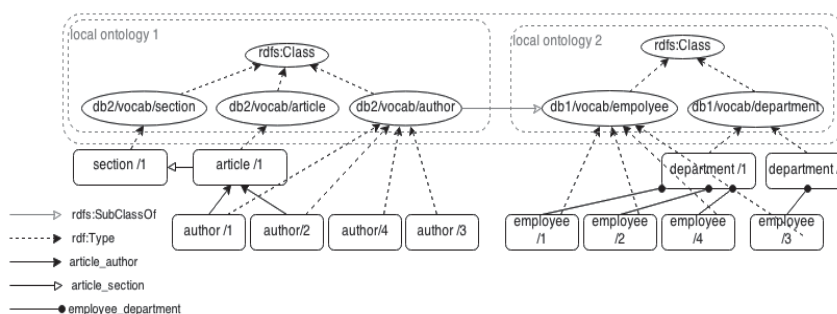


Рис. 5. Результат объединения источников

После объединения графов имеется возможность работать с ними как с единым целым при помощи языка SPARQL через модуль интеграции, либо непосредственно через интерфейс, который предоставляет сервер Fuseki. Так, например, мы можем вывести список всех сотрудников университета, которые имеют отношение к статье «article/1», выполнив следующий запрос:

```
«SELECT ?instance WHERE {?class rdfs:subClassOf <http://db1/vocab/employee>. ?instance rdf:type ?class. ?instance ?p <http://db2/article/1> }».
```

В результате выполнения данного запроса мы обнаружим, что сотрудниками университета, написавшими статью «article/1», являются «author/1» и «author/2». Таким образом, мы смогли извлечь информацию из двух несвязанных источников данных через единый словарь.

Как видно из примера, описанного выше, в запросе мы явно указываем, что нас интересуют лишь авторы – сотрудники университета. Таким образом, мы добавили семантической выразительности в запрос.

Информацию об авторах определенной статьи и при этом являющихся сотрудниками университета можно получить с помощью стандартного SQL. Для этого достаточно выполнить для первой БД следующий запрос: «SELECT last_name FROM author WHERE author_id in (SELECT author_id FROM article_author WHERE article_id=1)».

Затем для второй БД выполнить:

«*SELECT * FROM employee WHERE employee_name IN (результат выполнения предыдущего запроса)*».

Такой запрос для экспериментальных данных выполняется порядка 950 миллисекунд, но при этом тратится время на разработку кода, который выполняет извлечение и преобразование данных.

Данный запрос возможно выполнить и без написания вспомогательного кода, используя лишь стандартные возможности различных СУБД. Однако возможности СУБД имеют определенные ограничения. В случае если базы расположены на разных серверах, то применяются распределенные запросы (Oracle или MS SQL). Подобные подходы негативно сказываются на времени выполнения и гибкости SQL-запросов.

Одним из дополнительных преимуществ онтологического подхода является возможность использования механизмов построения логического вывода (reasoner). Применение подобного инструмента позволит автоматически дополнить результирующий граф возможными связями и отношениями, что в свою очередь позволит упростить процедуру построения SPARQL запросов и увеличить их гибкость. Применение данного механизма влечет за собой увеличение количества RDF троек. Так, после использования подобного механизма на экспериментальных данных общее число троек увеличилось до 42703 (на 17,84%).

4. Дальнейшие направления исследования

В рамках исследования, результаты которого представлены в данной статье, предполагаются следующие возможные направления:

1) Изучить и разработать методики для интеграции нереляционных источников данных. Например, можно провести аналогию между структурой данных RDF (субъект-предикат-объект) и структурой в документно-ориентированной БД (документ-ключ-значение).

2) Ввести понятия неопределенности/неточности для автоматизации процесса объединения онтологий и применить соответствующие математические модели.

3) Изучить и использовать технологии Linked Data для интеграции глобального информационного пространства с применением опубликованных словарей.

Заключение

В области обработки и управления информацией задача интеграции реляционных источников данных на сегодняшний день является актуальной. Одним из приоритетных направлений исследований в этой области считается применение онтологического подхода.

Онтологии применимы в системах интеграции данных, так как они предоставляют явную концептуализацию, которая описывается формализованным образом. Также языки описания онтологий обладают достаточной гибкостью и выразительностью для преодоления семантической гетерогенности.

Спроектированная и реализованная система интеграции, а также проведенные эксперименты доказывают состоятельность данного подхода и перспективность дальнейших исследований.

СПИСОК ЛИТЕРАТУРЫ

[Калиниченко, 2012] Калиниченко Л.А. Эффективная поддержка баз данных с онтологическими зависимостями: Реляционные языки вместо дескриптивных логик. Программирование, 2012. №6.

[Bertails et al., 2012] Bertails A., Arenas M., Prud'hommeaux E., Sequeda J., Editors. A Direct Mapping of Relational Data to RDF – <http://www.w3.org/TR/rdb-direct-mapping/>.

[Buccella et al., 2005] Buccella A., Cechich A. and Brisaboa N.R., Ontology-Based Data Integration Methods: A Framework for Comparison, Revista Colombiana de Computacion, 2005.

[Das et al., 2012] Das S., Sundara S., Cyganiak R., Editors. R2RML: RDB to RDF Mapping Language – <http://www.w3.org/TR/r2rml/>.

[Hert et al., 2011] Matthias H., Gerald R., Harald R. A Comparison of RDB-to-RDF Mapping Languages. In: Proceedings of the 7th International Conference on Semantic Systems (I-Semantics), Graz, Austria, 07 September 2011 – 09 September 2011.

[Prud'hommeaux et al., 2008] Prud'hommeaux E., Seaborne A., Editors. SPARQL Query Language for RDF – <http://www.w3.org/TR/rdf-sparql-query/>.

[Villazón-Terrazas et al., 2012] Villazón-Terrazas B., Hausenblas M. RD-B2RDF Implementation Report – <http://www.w3.org/TR/rdb2rdf-implementations/>.

[Wache et al., 2001] Wache H., Vögele T., Visser U., Stuckenschmidt H., Schuster G., Neumann H., and Hübner S. Ontology-Based Integration of Information – A Survey of Existing Approaches. In Proceedings of the IJCAI-01 Workshop on Ontologies and Information Sharing, 2001.

УДК 519.17

ЗАДАЧИ И МЕТОДЫ АНАЛИЗА СХОДСТВА ТЕМПОРАЛЬНЫХ ОРГРАФОВ

В.В. Кохов (*viktor-kokhov@ya.ru*)

Национальный исследовательский университет
Московский энергетический институт, Москва

С целью исследования динамики изменения структурных свойств темпоральных орграфов выделены два класса задач определения сходства темпоральных орграфов. Предложены два подхода к разработке методов решения задач. Приведен сравнительный анализ методов из двух предложенных подходов.

Введение

Как выделено в [Айзерман, 1977], [Bramsens, 2006], [Берштейн и др., 2010], в настоящее время наиболее актуальным направлением исследований в области структурного анализа систем является разработка методов анализа графов с изменяемой структурой во времени (темпоральных орграфов (T -орграфов)). Работа [Айзерман, 1977] была началом исследований по графодинамике, т.е. по динамическому описанию структур. В ней выделены базовые классы задач и в качестве наиболее значимых в теоретическом и прикладном аспектах выделены: 1) *задача определения расстояния*, которое помогает ввести представление об устойчивости изменений структуры графа во времени (графовых траекторий) по отношению к малым возмущениям и о монотонности в смысле этого расстояния процессов в графодинамике; 2) *задача определения сходства-различия между двумя графами*.

Примерами прикладных задач являются анализ изменений: 1) административных структур, обычно описываемых орграфами-деревьями; 2) организации систем связи, снабжения; 3) структур корпоративных социальных сетей, сетей коммуникаций сотрудников фирм с целью анализа и принятия обоснованных управленческих решений; 4) структур динамических семантических сетей и др. Ниже с целью определения монотонности (немонотонности) процесса изменения структур T -орграфов выделяются два клас-

са задач анализа сходства: 1) сходство структур одного T -орграфа; 2) сходство структур двух T -орграфов.

В работе применяется терминология, используемая в [Bramsen, 2006], [Берштейн и др., 2010]. Предлагаются два подхода к разработке методов анализа сходства: 1) структурно-характеристический подход (СХ-подход); 2) подструктурно-метрический подход (ПМ-подход). СХ-подход основан на вычислении сходства-расстояния между парой T -орграфов с использованием индексов и вектор-индексов сложности, вычисленных в расширяемом базисе структурных дескрипторов. ПМ-подход использует методы определения максимальных общих фрагментов двух T -орграфов или их структурных инвариантов. В качестве одной из центральных теоретических проблем выделена проблема определения необходимых и/или достаточных условий, накладываемых на структуры двух T -орграфов, приводящих к выполнению условия монотонности изменения функции сходства на анализируемых дискретах времени в заданном промежутке.

1. Основные определения

T -орграфом назовем тройку $G=(V^{(t)},E^{(t)},T)$, где $V^{(t)}$ – множество вершин в момент времени t с числом вершин $|V^{(t)}|=p$, $T=\{t_1,t_2,\dots,t_j,\dots,t_n\}$ – множество натуральных чисел, определяющих (дискретное) время, $E^{(t)}=\{\Gamma_i\}$ – семейство соответствий или отображений множества вершин $V^{(t)}$ в себя в момент времени $t \in T$, т.е. $(\forall t \in T) \Gamma_i: V^{(t)} \rightarrow V^{(t)}$. Через tG обозначим T -орграф в момент времени t . Фрагмент f орграфа tG в момент времени t получается при удалении дуг или вершин и дуг. Орграф $t_1G=(V^{(t_1)},E^{(t_1)},T)$ изоморфен орграфу $t_2G=(V^{(t_2)},E^{(t_2)},T)$ ($t_1G \approx t_2G$), если

$$\exists \varphi: (V^{(t_1)} \leftrightarrow V^{(t_2)}) \& (\forall v_i, v_j \in V^{(t_1)} [(v_i, v_j) \in E^{(t_1)} \leftrightarrow (\varphi(v_i), \varphi(v_j)) \in E^{(t_2)}]),$$

где $\varphi(v_i), \varphi(v_j) \in V^{(t_2)}$. Орграф $t_1G=(V^{(t_1)},E^{(t_1)},T)$ изоморфно вкладывается в орграф $t_2G=(V^{(t_2)},E^{(t_2)},T)$ ($t_1G \subseteq^f t_2G$), если в t_2G есть фрагмент $f = t_2G_*=(V_*^{(t_2)},E_*^{(t_2)},T)$, для которого выполняется условие $t_1G_* \approx t_2G$.

Множество всех изоморфизмов орграфа tG на себя образует группу по умножению подстановок φ и обозначается через $Aut(tG)$, порядок группы – через $|Aut(tG)|$. Под числом *канонических изоморфных вложений* tG_* в tG будем понимать величину, определяемую следующим образом: $w(tG_*,tG)=W(tG_*,tG)/|Aut(tG_*)|$, где

$W(tG^*, tG)$ – число всех изоморфных вложений tG^* в tG . Под максимальным общим фрагментом двух орграфов t_1G, t_2G ($MCF(t_1G, t_2G)$) понимаем орграф tG^* , для которого справедливы условия: а) $tG^* \subseteq t_1G$ и $tG^* \subseteq t_2G$; б) не существует большего tG^* по числу дуг фрагмента в t_1G , для которого выполняется условие а).

Основным инструментом при вычислении характеристик tG является использование инвариантов tG (инвариантов, характеризующих расположение фрагментов в tG). Обозначим множество всех T -орграфов через \mathfrak{R} . Пусть R – отношение «быть изоморфными» орграфами tG , а Q – непустое множество с отношением эквивалентности τ (множество чисел, векторов, матриц и т.д.). Функция IN , заданная на множестве \mathfrak{R} и принимающая значения в Q , называется *инвариантом* tG , если справедливо условие

$$\forall t_iG, t_jG \in \mathfrak{R} [t_iG(R)t_jG \rightarrow IN(t_iG)(\tau)IN(t_jG)].$$

Пусть $F^{(lk)}(tG) = \{t_1^{(lk)}, t_2^{(lk)}, \dots, t_m^{(lk)}\}$ обозначает множество помеченных фрагментов типа k в tG . Инвариантом, характеризующим расположение фрагмента tf^{lk} , называется функция $in(tf^{lk})$, заданная на множестве $F^{lk}(tG)$, принимающая значения в Q и удовлетворяющая условию

$$\forall t_i^{(lk)}, t_j^{(lk)} \in tF^{(lk)} [t_i^{(lk)}(\xi)t_j^{(lk)} \rightarrow in(t_i^{(lk)})(\tau)in(t_j^{(lk)})],$$

где ξ^t – отношение эквивалентности «принадлежать одной и той же орбите k -группы». Аналогично понятиям, связанным с инвариантом орграфа tG , вводятся понятия для инварианта $in(tf^{lk})$.

2. Постановка задач определения сходства-различия для структур T -орграфа

Выделим два класса задач *определения сходства-различия*: 1) *структур одного T -орграфа*; 2) *структур двух T -орграфов*. Пусть $D(t_iG_k, t_jG_l)$ – метрика (псевдометрика) или функция для вычисления расстояния между структурами t_iG_k, t_jG_l , где $t_p, t_j \in T$. В первый класс входят две следующие задачи.

Задача 1.1. Заданы $G=(V^{(t)}, E^{(t)}, T)$, где $T=\{t_1, t_2, \dots, t_p, \dots, t_n\}$ и метрика $D(t_iG_k, t_jG_l)=D(t_iG, t_jG)$.

Необходимо для $\forall(t_p, t_{i+1})$, где $i=1..(n-1)$, определить $D(t_iG, t_{i+1}G)$ и построить функцию fd_1 изменения значений вычисленных расстояний.

Задача 1.2. Заданы $G=(V^{(t)},E^{(t)},T)$, где $T=\{t_1,t_2,\dots,t_j,\dots,t_n\}$ и метрика $D(t_iG_k,t_jG_l)=D(t_iG,t_jG)$.

Необходимо для $\forall(t_i,t_j)$, где $j=2..n$, определить $D(t_1G,t_jG)$ и построить функцию fd_2 изменения значений вычисленных расстояний.

Постановки задач сравнительного анализа динамики изменения сходства двух T -орграфов для второго класса имеет следующий вид:

Задача 2.1. Заданы $G_1=(V_1^{(t)},E_1^{(t)},T)$, $G_2=(V_2^{(t)},E_2^{(t)},T)$, где $T=\{t_1,t_2,\dots,t_j,\dots,t_n\}$ и метрика $D(t_iG_k,t_jG_l)=D(t_iG_1,t_jG_2)$.

Необходимо для $i=1..n$ определить $D(t_iG_1,t_iG_2)$ и построить функцию fd_1 изменения значений вычисленных расстояний.

Задача 2.2. Заданы $G_1=(V_1^{(t)},E_1^{(t)},T)$, $G_2=(V_2^{(t)},E_2^{(t)},T)$, где $T=\{t_1,t_2,\dots,t_j,\dots,t_n\}$ и метрика $D(t_iG_k,t_jG_l)=D(t_iG_1,t_jG_2)$.

Необходимо для $j=2..n$ определить $\Delta = D(t_iG_1,t_jG_2) - D(t_lG_1,t_jG_2)$ и построить функцию fd_2 изменения значений Δ .

В качестве задачи третьего подкласса, определяющей интегральное сходство-различие двух T -орграфов, сформулируем следующую задачу.

Задача 2.3. Заданы $G_1=(V_1^{(t)},E_1^{(t)},T)$, $G_2=(V_2^{(t)},E_2^{(t)},T)$, где $T=\{t_1,t_2,\dots,t_j,\dots,t_n\}$ и задана метрика для вычисления расстояния:

$$D(G_1,G_2)=D(MCF(t_1(G_1),t_2(G_1),\dots,t_n(G_1)),MCF(t_1(G_2),t_2(G_2),\dots,t_n(G_2)))$$

Необходимо вычислить значение $D(G_1,G_2)$.

3. Структурно-характеристический подход к анализу сходства T -орграфов

Основу СХ-подхода для решения этих задач определяет использование двух видов инвариантов tG : 1) индекс сложности; 2) вектор-индекс сложности [Kokhov, 2012]. Индексы и вектор-индексы tG характеризуют количественный и качественный состав tG в заданном базисе фрагментов. Под *структурным спектром* $SS(tG/B)$ орграфа tG в базисе $B=\langle b_1,b_2,\dots,b_i,\dots,b_{k1} \rangle$ будем понимать запись следующего вида: $SS(tG/B)=\langle \alpha_1,\alpha_2,\dots,\alpha_i,\dots,\alpha_{k1} \rangle$, где $\alpha_i=1$, если $b_i \in B$ изоморфно вкладывается в tG и $\alpha_i=0$ – в противном случае. Пусть для tG построен его *полный структурный спектр* (ПСС) в базисе B :

$$FSS(tG/B)=\langle w(tG/b_1),w(tG/b_2),\dots,w(tG/b_i),\dots,w(tG/b_{k1}) \rangle,$$

где b_i – фрагмент базиса; $w(b_i)$ – число канонических изоморфных вложений фрагмента b_i в tG ; k_1 – число фрагментов базиса B , относительно которого характеризуется сложность tG . Примем $ISC(tG=K_1)=1$, $ISC(tG=K_2)=3$. Так как для любого фрагмента f_i орграфа tG можно определить его ПСС, а для каждого фрагмента от фрагмента tG можно построить его ПСС и т.д., то рекурсивным методом всегда можно вычислить индекс сложности (ISC) и вектор-индекс сложности (V_ISC) для tG в базисе B :

$$ISC(tG/B)=w(b_1)ISC(b_1)+\dots+w(b_{k_1})ISC(b_{k_1}),$$

$$V_ISC(tG/B)=\langle w(b_1)ISC(b_1); \dots; w(b_{k_1})ISC(b_{k_1}) \rangle.$$

Через $V_RSC(tG/B)$ обозначим вектор-индекс *относительной спектральной сложности* орграфа tG в базисе B , где

$$V_RSC(tG/B)=\langle w(b_1)ISC(b_1)/ISC(tG); \dots; w(b_{k_1})ISC(b_{k_1})/ISC(tG) \rangle.$$

Вычисление значений $V_RSC(tG/B)$ позволяет определять суммарную относительную значимость фрагментов $b_i \in B$ в общей сложности tG . Заметим, что задача определения всех простых путей *всех возможных длин* между парой вершин орграфа является NP -полной проблемой [Гэри и др., 1982]. По разнице значений индексов SC -сложности легко определить количественную меру сходства t_iG_k и t_jG_b , проводить мониторинг изменения структур двух T -орграфов для заданного промежутка времени и определять тенденции изменения сложности структур. Аналогичные исследования будем проводить с использованием вектор-индексов и евклидовой метрики. Использование расширяемых базисов B , например путей, полупутей, ордеревьев и т.д., позволит проводить все более и более точный анализ сходства.

4. Подструктурно-метрический подход к анализу сходства T -орграфов

Расширим концепцию классического ПМ-подхода [Bunke H. et al., 1998] на класс орграфов и T -орграфов. Пусть $MCS(t_1G, t_2G)$ обозначает максимальный общий *подграф* для t_1G и t_2G . Тогда коэффициент сходства SI , несходства DSI и расстояние D между t_1G и t_2G , будем вычислять по формулам:

$$1. D_1(t_1G_1, t_2G_2)=p(t_1G_1)+p(t_2G_2)-2p(MCS(t_1G_1, t_2G_2));$$

2. $D_2(t_1G_1, G_2) = p(t_1G_1) + q(t_1G_1) + p(t_2G_2) + q(t_2G_2) - 2(p(MCF(t_1G_1, t_2G_2)) + q(MCF(t_1G_1, t_2G_2)))$;
3. $SI_1(t_1G_1, G_2) = p(MCS(t_1G_1, t_2G_2))^2 / (p(t_1G_1)p(t_2G_2))$;
4. $DSI_1(t_1G_1, G_2) = 1 - SI_1(t_1G_1, t_2G_2)$;
5. $SI_2(t_1G_1, t_2G_2) = \frac{(p(MCF(t_1G_1, t_2G_2)) + q(MCF(t_1G_1, t_2G_2)))^2}{(p(t_1G_1) + q(t_1G_1))(p(t_2G_2) + q(t_2G_2))}$;
6. $DSI_2(t_1G_1, G_2) = 1 - SI_2(t_1G_1, t_2G_2)$.

Заметим, что задача определения MCF или MCS двух орграфов принадлежит классу NP [Гэри и др., 1982]. Новым аспектом развития ПМ-подхода к анализу сходства T -орграфов является использование MCF и MCS , вычисляемых для всех изменений структуры как первого, так и второго из T -орграфов в заданных промежутках времени.

5. Примеры решения задач разными методами

Пусть структуры t_1G - t_5G представляют T -орграф G в моменты времени t_1 - t_5 (рис. 1).

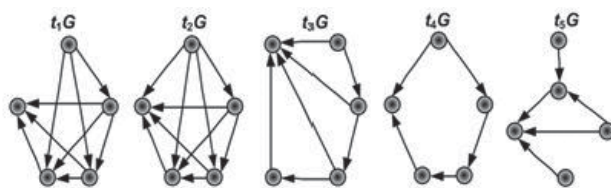


Рис. 1

В табл. 1 приведены значения индексов СС-сложности в базисах путей P и полупутей HP , значения попарных расстояний ($D_2(t_iG, t_jG)$) (выше главной диагонали) и индексов сходства ($SI_2(t_iG, t_jG)$) (ниже главной диагонали) для T -орграфа (рис. 1).

Таблица 1

t	$ISC(tG/P)$	$ISC(tG/HP)$	tG	t_1G	t_2G	t_3G	t_4G	t_5G
t_1	383	2356	t_1G	0	2	4	6	4
t_2	334	2284	t_2G	0,86	0	2	4	4
t_3	270	959	t_3G	0,72	0,86	0	2	4
t_4	215	474	t_4G	0,58	0,71	0,83	0	4
t_5	38	192	t_5G	0,58	0,71	0,68	0,64	0

На рис. 2 показаны графики, характеризующие динамику изменения значений индексов сложности для T -орграфа (рис. 1).

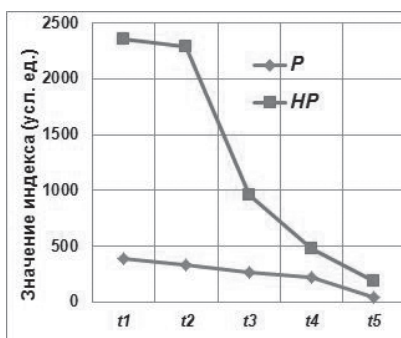


Рис. 2

На рис. 3–4 приведены результаты решения с использованием СХ-подхода задачи 1.1, а на рис. 5–6 – вида 1.2. Результаты, приведенные на рис. 3–6, приводят к выводам о монотонном характере динамики изменения сходства при использовании СХ-подхода при решении задачи 1.2 и немонотонном – при решении задачи 1.1. Заметим, что базис полупутей является расширением базиса путей и более точно характеризует разветвленность структуры $t_i G$.

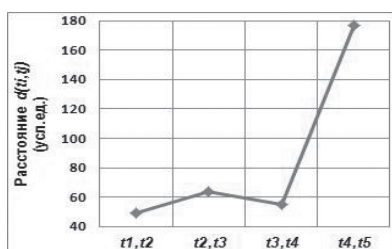


Рис. 3

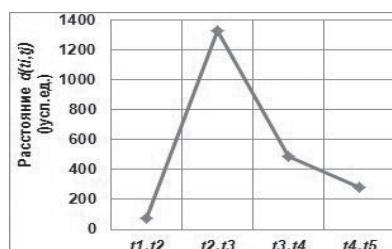


Рис. 4

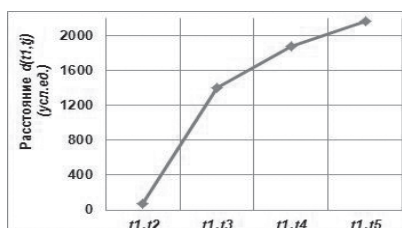


Рис. 5

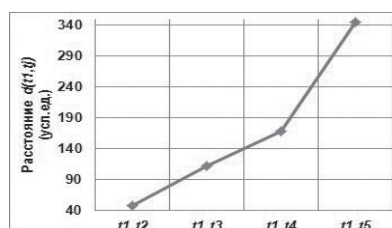


Рис. 6

На рис. 7–8 приведены результаты решения задач 2.1 и 2.2 с использованием ПМ-подхода. Динамика изменения сходства для анализируемого примера с использованием D_2 и S_2 имеет монотонный характер при решении задачи 2.1 и немонотонный – при решении задачи 2.2.

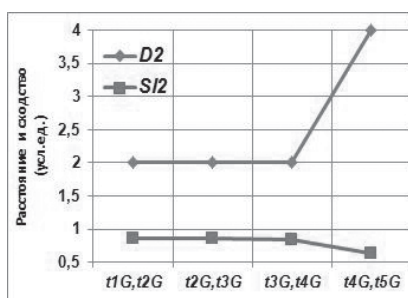


Рис. 7

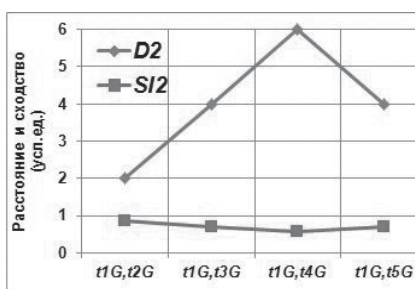


Рис. 8

В основе решения задачи 2.3 лежит определение максимальных стационарных фрагментов в структурах каждого из T -орграфов. На рис. 9 приведены $MCF(G_1)$, $MCF(G_2)$ и $MCF(MCF(G_1), MCF(G_2)) = MCF(G_1, G_2)$. Результатами решения задачи 2.3 являются значения расстояния и сходства:

$$D_2(G_1, G_2) = 5 + 9 + 6 + 10 - 2(5 + 9) = 2,$$

$$S_2(G_1, G_2) = (5 + 8)^2 / ((5 + 9)(6 + 10)) = 0,75.$$

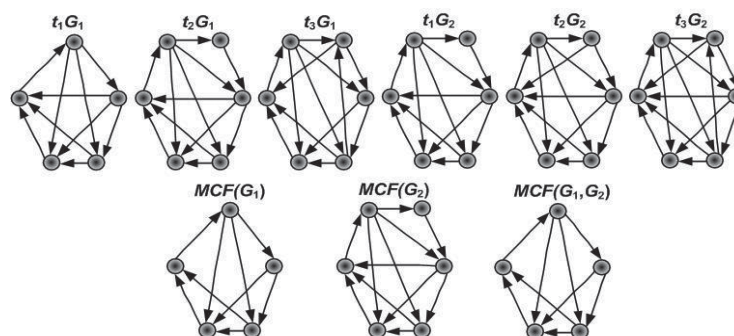


Рис. 9

Заключение

Результаты объемных вычислительных экспериментов показали более низкий уровень чувствительности и существенно большую вычислительную сложность при решении задач сходства T -орграфов по ПМ-подходу в сравнении с использованием СХ-подхода. При ограничении длины базиса методы СХ-подхода имеют полиномиальную вычислительную сложность при достаточно высокой точности характеристики структур T -орграфов в базисе путей или полупутей. Однако методы ПМ-подхода дают возможность выделять стационарные фрагменты максимального размера и решать оригинальный класс задач сравнения двух T -орграфов на основе сходства их стационарных подструктур. Предложенные методы программно реализованы и используются в учебном процессе НИУ МЭИ, НИУ ВШЭ, научных исследованиях динамических сетей в ИВМиМГ СО РАН.

СПИСОК ЛИТЕРАТУРЫ

[Айзерман и др., 1977] Айзерман М.А., Гусев Л.А., Петров С.В., Смирнова. Динамический подход к анализу структур, описываемых графами (основы графодинамики) // *АиТ*. – 1977. – №7.

[Берштейн и др., 2010] Берштейн Л.С., Боженюк А.В. Использование темпоральных графов как моделей сложных систем // *Известия ЮФУ. Технические науки*. 2010. – № 4 (105).

[Гэри и др., 1982] Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи: Пер. с англ. – М.: Мир. 1982.

[Bramsen, 2006] Bramsen P.J. Doing Time: Inducing Temporal Graphs. Technical report, Massachusetts Institute of Technology. 2006.

[Bunke H. et al., 1998] Bunke H., Sharer K. A Graph Distance Metric Based on the Maximum Common Subgraph // *Pattern Recognition Letters*, vol. 19, no. 3–4, 1998, pp. 255–259.

[Kokhov, 2012] Kokhov V.A. Two Approaches to Determining Similarity of Two Digraphs // *Journal of Computer and Systems Sciences International*, 2012, Vol. 51, pp. 695–714.

УДК 004.822

**ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ
ДЛЯ ЛИНГВИСТИЧЕСКИ-ОРИЕНТИРОВАННОГО
ПОПОЛНЕНИЯ ОНТОЛОГИИ¹**

С.С. Курбатов (*curbatow.serg@yandex.ru*)

А.П. Лобзин (*lobzin@nicevt.ru*)

Г.К. Хахалин (*khakhalin@yandex.ru*)

Научно-исследовательский центр электронной вычислительной
техники, Москва

Описана логика создания программного обеспечения, ориентированного на пополнение онтологии с использованием средств лингвистической трансляции. Предложен интерфейс с онтологией, обеспечивающий ввод соответствия естественно-языковых описаний и формальных конструкций. Разработан механизм выдвижения предположений по обобщению вводимой информации с использованием онтологии и по организации проверки предположений. Реализован макет, воплощающий разработанную логику в программах.

Введение

Проблема автоматизации построения и модификации онтологий сохраняет свою актуальность практически с момента появления онтологического подхода, развивающего концепции традиционных систем представления знаний (knowledge representation). Ряд широко известных систем создания и редактирования онтологий позволяет достаточно быстро реализовать специализированную онтологию для некоторой предметной области [Гаврилова, 2006]. Однако для масштабной онтологии в реальной предметной области возникает проблема автоматизации рутинных процессов при заполнении онтологии. Проблема усугубляется потребностями обучения новых пользователей и разработки четкой и легко понимаемой документации [Smith, 2006]. В данной работе эта проблема исследована в аспекте отображения естественно-языковых описаний в алгоритмически-программные концепты с акцентом на механизм автоматизированного пополнения онтологии, содержащей эти концепты.

¹ Работа выполнена при финансовой поддержке РФФИ (проект 12-07-00531а).

1. Общая логика

Основная логика предполагает интерактивное взаимодействие с онтологией и обмен в общем случае двумя различными типами информации: естественно-языковое описание и формальная конструкция языка программирования. В качестве формальной конструкции может выступать текст макросов EXCEL, структурного описания графического объекта (язык *grasp*, раздел 3), запрос на SQL, текст на языке 4L (для СУБД Progress) и т.п.

Интерпретация ЕЯ-описания в онтологии позволяет не только зафиксировать соответствие ЕЯ-описания и формализма, но и выдвинуть предположения по обобщению введенной информации. В процессе дальнейшего взаимодействия эти предположения могут быть подтверждены, опровергнуты или уточнены. Интерпретация предполагает использование эвристических правил, относящихся как к ЕЯ-описаниям, так и к формальным конструкциям.

Использовалась онтология, реализованная на СУБД Progress. Онтология базируется на структурированной семантической сети и включает лингвистические знания (в частности, правила перифразирования), а также концепты соответствий ЕЯ-описаний и формальных структур. Ряд формальных моментов, относящихся к онтологии, приведен в [Курбатов, 2014].

2. ЕЯ-описания графических фигур EXCEL и макросы

В данном разделе описывается отработка логики лингвистически-ориентированного пополнения онтологии с использованием языка макрокоманд для базовых графических фигур EXCEL. Основной цикл взаимодействия включал создание графического элемента или выполнение действия над ним и передачу онтологии ЕЯ-описания операции и формального текста на языке макросов EXCEL. При этом операция создания в основном сводилась к фиксации в онтологии соответствия ЕЯ и формального описаний, а для действия в онтологии выдвигались предположения о возможном обобщении введенных структур.

Проиллюстрируем взаимодействие на примере. Средствами EXCEL рисовалась окружность и вызывалась программа интерфейса с онтологией (в дальнейшем для краткости – онтология). Передаваемые параметры, в частности, имели такой вид:

«это окружность» и «ActiveSheet.Shapes.AddShape (9, 100, 107, 50, 50)». Затем окружность в EXCEL перемещалась вниз и передаваемые параметры получали вид «окружность сдвинута вниз» и «ActiveSheet.Shapes.AddShape (9, 100, 269, 50, 50)». На основании такой информации программа интерфейса с онтологией выдвигала предположение, что именно увеличение числа 107 до 269 приводит к выполнению действия «сдвинуть».

Более детально – вызывается онтологическая функция сравнения символьных строк, определяется первое несовпадение и вызывается онтологическая функция выделения целого числа в символьной строке. Далее после сравнения чисел 107 и 269 выдвигается предположение, что увеличение целого числа в соответствующей позиции соответствует действию «сдвинуть». Важно, что действие описано качественно («увеличить»), а не количественно (конкретным числом).

Наличие в онтологии такой информации позволяет для произвольной окружности в EXCEL на ввод фразы «сдвинуть окружность вниз» выполнить действие в соответствии с предположением. Важно также, что при наличии в лингвистической модели соответствующих знаний естественно предположение о сдвиге «вверх» как антониме «вниз». Разумеется, модель должна также содержать знания об «уменьшить» как антониме «увеличить» и его связи с соответствующим качественным представлением в онтологии. Наличие в онтологии общего концепта типа «перемещение» позволяет также сформулировать пользователю запросы, содержащие антонимы «вправо», «влево».

Более сложный пример относится к описанию расположения отрезков, образующих угол или соприкасающихся отрезков. В этом случае необходимо определить совпадение двух чисел, определяющих начала или концы отрезков. Поэтому после неудачи с выделением одного числа вызывается программа выделения последовательности целых чисел для произвольной символьной строки. Далее выполняется программа выделения совпадающих элементов для двух произвольных последовательностей, и онтология может выдвинуть предположение о неслучайном совпадении этих элементов, описывающих именно отношение «соприкасаются» для отрезков.

Дальнейшее расширение описанной техники связано с описанием таких отношений между отрезками, как «вертикальны», «го-

горизонтальны», «параллельны», «пересекаются»; вводом фигуры «овал» и т.д. Качественное расширение предполагает развитие программные средства, позволяющие формировать в онтологии древовидные структуры, соответствующие формальным структурам произвольной вложенности. Предполагается также, что для элементов алгоритма разбора такой структуры будет дано ЕЯ-описание и соответствующая формальная структура.

Ряд ad hoc вариантов был реализован, типичным является ЕЯ-описание «найти выражение, сбалансированное по круглым скобкам и не содержащее круглые скобки внутри» и соответствующий шаблон для поиска регулярного выражения – $\backslash([\wedge\wedge])^*\backslash$. Лингвистическая трансляция ЕЯ-описаний шаблонов для регулярных выражений с использованием метапеременных онтологического уровня является предметом специального исследования, ряд аспектов этой проблемы для русского языка отражен в [Мерзляков, 2010], а для английского языка в [Kushman et al., 2013]. Проводимые нами исследования ориентированы на русский язык с возможностью перефразирования ЕЯ-описания поискового запроса и получения канонической структуры шаблона для поиска регулярного выражения.

3. Язык описания графических объектов GRASP

Более интересные графические объекты, чем базовые фигуры EXCEL (или WORD), можно описать на языке grasp [Литвинович, 2012]. Этот язык позволяет на уровне объектов, их свойств и базового отношения сцепки описать достаточно интересные объекты. Высокий уровень языка позволил организовать его эффективный интерфейс с онтологией и возможность синтеза графических объектов по описаниям на ЕЯ. С использованием возможностей языка grasp, прикладной онтологии и средств упрощенной лингвистической трансляции была реализована система класса ТТР (Text-To-Picture) [Литвинович, 2013], наиболее интересные в аспекте целей данной статьи, возможности которой описываются ниже.

3.1. ЕЯ-описания графических образов на grasp

Синтез графических объектов по тексту на ЕЯ был намечен еще в пионерских работах [Литвинцева и др., 1991] и был ориентирован на задачи интеллектуальных САПР. В вышеупомянутой

ТТР-системе акцент сделан на задачи генерации по ЕЯ-описанию программного текста на grasp, в результате интерпретации которого синтезируются графические объекты достаточно сложной структуры.

Синтез выполняется упрощенной лингвистической трансляцией ЕЯ-описания графического объекта в структуры прикладной онтологии, генерацией из этой структуры текста на grasp и выполнением текста интерпретатором grasp для получения изображения средствами OpenGL.

Ниже дан пример ЕЯ-описания (курсивом), синтезированный графический объект выводится в трехмерном изображении.

В точке 0 арматуры присоединена цистерна Tank_B через точку 0.

В точке 1 арматуры присоединена другая цистерна Tank_B через точку 0.

В точке 8 арматуры присоединен адаптер через точку 0.

В точке 1 адаптера присоединена другая арматура через точку 0.

В точке 5 этой арматуры присоединена солнечная панель типа С через точку 0.

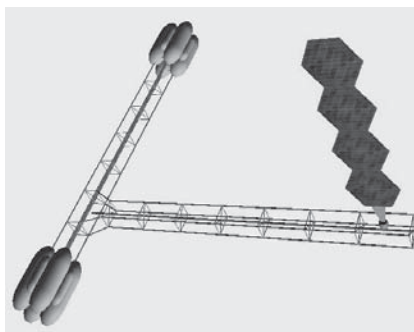


Рис. 1. Синтезированный по ЕЯ-описанию графический объект

Трансляция ЕЯ-описания использует адаптированный для целей ТТР-системы механизм правил перефразирования, позволяющий свести входные предложения к каноническим структурам. Важно отметить наличие в системе средства лингвистической адаптации, допускающего ввод на ЕЯ нового правила перефразировки. При

этом конкретное правило, введенное пользователем, обобщается с учетом информации из онтологии.

Поясним возможности адаптации на примере. Пусть пользователь задает конкретную фразу «1 и 2 являются точками сцепки арматуры и цистерны» и сообщает системе, что эта фраза эквивалентна следующей «1 и 2 являются точками сцепки арматуры и цистерны». При этом вторая фраза сводится к каноническому виду и для нее синтезируется графический объект. В системе при этом формируется обобщенное правило, левая часть которого имеет вид – {N1} и {N2} являются точками сцепки {OB1} и {OB2}.

В общем случае синтезируемый графический объект описывается древовидной структурой *grasp*, вершины которой соответствуют базовым блокам. Свойства блоков задают специфику типа цвет, размер, текстуру, особые характеристики (например, число секций) и т.п. Графический объект выводится по умолчанию как 3D, возможно отображение как 2D. После трансляции ЕЯ-описания объекта в онтологическое представление выполняется генерация программного текста на языке *grasp* и вывод графического изображения.

3.2. Скрипты *grasp* и ЕЯ-описания

Заполнение онтологии для экспериментов с *grasp* выполнялось с помощью достаточно простых инструментальных средств. Однако для задач большего масштаба необходимы более развитые средства. С целью разработки таких средств были проведены эксперименты по автоматической генерации текстов на *grasp*, описывающих базовые объекты, и выводу изображения администратору онтологии для ЕЯ-описания объекта. Большим преимуществом такого подхода является то, что администратор имеет дело только с ЕЯ-описанием и графическим изображением, он изолирован от текста на *grasp*. Разумеется, обращение к уровню *grasp* возможно, но оно потребует только при возникновении особых ситуаций (ошибок, неоднозначностей и т.п.).

Полноценный вариант для такого рода экспериментов предполагает лингвистическую трансляцию документации по *grasp* и выдвижение предположения по формальному описанию базовых объектов. Однако этот вариант является предметом специального исследования, тем не менее, в процессе экспериментов мы стремились по возможности учитывать действия человека, работающего с документацией по *grasp*.

Разработка программных средств включала:

- упрощенный анализ описания grasp на ЕЯ;
- программу просмотра файлов из оглавления blocks (по результатам анализа);
- генерацию текстов grasp для базовых объектов и вывод изображений;
- фиксацию в онтологии ЕЯ-описаний и текстов grasp.

В целом эксперименты показали работоспособность разработанных средств и перспективность их использования для пополнения онтологии лингвистически-ориентированными методами.

4. ЕЯ-описания объектов базы данных и программ интерфейса к базе

Более масштабное использование лингвистических средств для автоматизации пополнения онтологии описано в [Курбатов и др., 2013]. В этом случае онтология содержит информацию о прикладной области, схеме и содержании базы данных (информационной модели прикладной области), и о программном обеспечении для доступа к базе данных. База данных, программы интерфейса и документация на систему доступа к базе выступают как исходные данные для заполнения онтологии.

Программы интерфейса и схема базы задают обобщенную информацию о предметной области (сеть торговых предприятий) и образуют оболочку (систему), заполняемую конкретной информацией. Система определяет базовую бизнес-логику, но специфика применения определяется при заполнении оболочки:

- наименования объектов (товаров, предприятий);
- классификация объектов;
- наименования свойств объектов.

Наиболее важной характеристикой предлагаемой онтологии, описывающей такую систему, является существенное использование профессионально-ориентированного естественного языка для широкого спектра объектов: наименования таблиц и полей базы, а также значений полей; наименования программных модулей системы и ЕЯ-описания выполняемых ими функций; наименование и назначение параметров процедур. Наиболее важные цели такой онтологии:

- ЕЯ-доступ к содержимому базы данных и программам интерфейса;
- генерация программ модификации интерфейса по ЕЯ-описанию.

Инструментальные средства, не использующие лингвистические методы, позволили ввести в онтологию описания таблиц/полей и тексты программ интерфейса. Путем просмотра текстов схемы базы данных и программ интерфейса формировалось онтологическое описание объектов типа таблицы, полей и программных модулей. Между этими объектами были построены отношения типа *<является полем таблицы>*, *<вызывает модуль>*, *<используется в модуле>* и т.п.

Далее онтология была дополнена соответствиями ЕЯ-описания функционала системы (в основном подготовленного вручную) и некоторого модуля интерфейса. Соответствие выявлялось с помощью программных средств, анализирующих тексты интерфейса на предмет ЕЯ-наименований пунктов меню, сопоставляющихся с описанием функционала. Например, функционал с ЕЯ-описанием «отчет об остатках», выявленный на этапе лингвистической структуризации и полностью совпадающий с пунктом меню, был сопоставлен с модулем, вызываемым из этого пункта.

В процессе экспериментов в онтологию были занесены десятки ЕЯ-описаний, соответствующих базовому описанию функционала из документации на систему. Кроме того, из описателей DESCRIPTION полей и таблиц схемы базы данных в онтологию заносились ЕЯ-описания этих объектов. Дополнительно тексты программ интерфейса сканировались на ЕЯ-описания, содержащиеся в примечаниях. Если содержание примечаний удавалось интерпретировать как ЕЯ-описание, то оно также заносилось в онтологию и ассоциировалось с соответствующим модулем интерфейса.

Заключение

Совершенствование механизмов пополнения онтологии ЕЯ-описаниями объектов базы данных и программ интерфейса к этой базе является основным направлением развития макетного варианта разработанного ПО. Важная прикладная задача такого развития – синтез программ для работы с базой данных по ЕЯ-описаниям. Эксперименты с фигурами EXCEL, с графическими описаниями на

grasp и т.п. выполнялись в основном для отработки общей методики генерации программного кода для произвольного языка программирования (макросы EXCEL, SQL, 4GL, grasp и т.д.). Элементы такой методики нашли воплощение в рамках реализованного программного макета.

Благодарим разработчиков выбранной для экспериментального исследования системы управления сетевой торговлей за консультации и ценные советы по улучшению интерфейса с онтологией.

СПИСОК ЛИТЕРАТУРЫ

[**Kushman et al., 2013**] Kushman Nate, Barzilay Regina. Using Semantic Unification to Generate Regular Expressions from Natural Language // Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology – 2013.

[**Smith, 2006**] Smith Barry. Against Idiosyncrasy in Ontology Development // in B. Bennett and C. Fellbaum (Eds.), Formal Ontology and Information Systems, (FOIS 2006), Baltimore November 9–11 2006.

[**Гаврилова, 2006**] Гаврилова Т.А. Формирование прикладных онтологий // Труды X национальной конференции по искусственному интеллекту с международным участием – КИИ-2006. М.: Физматлит, 2006.

[**Курбатов и др., 2013**] Курбатов С.С., Лобзин А.П., Хахалин Г.К. Онтология для автоматизированного синтеза программ по описанию на предметно-ориентированном естественном языке // Труды IV Всероссийской конференции с международным участием «Знания – Онтологии – Теории» (ЗОНТ-13). – Новосибирск: Институт математики СО РАН, 2013.

[**Литвинович, 2012**] Литвинович А.В. Язык описания графических объектов GRASP // Нейрокомпьютеры: разработка, применение, № 10, 2012.

[**Литвинович, 2013**] Литвинович А.В. Система синтеза изображений по тексту на естественном языке // Динамика сложных систем – XXI век, № 1, 2013.

[**Литвинцева и др., 1991**] Литвинцева Л.В., Поспелов Д.А. Визуализация пространственных сцен по текстовым описаниям для интеллектуальных систем // Известия АН СССР, Техническая кибернетика, №5, 1991.

[**Мерзляков, 2010**] Мерзляков Д.А. Генерация регулярных выражений для автоматизации проверки текстов открытого характера. Пермский государственный национальный исследовательский университет, 2010.

[**Курбатов, 2014**] Онтология, <http://eia--dostup.ru/Онтология.htm>.

УДК 004.82

**ОСОБЕННОСТИ ИНТЕГРАЦИИ МЕХАНИЗМОВ
ЛОГИЧЕСКОГО ВЫВОДА В ОНТОЛОГИЧЕСКУЮ
МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ С ПОМОЩЬЮ
SWRL-ПРАВИЛ**

В.С. Мошкин (*postforvadim@yandex.ru*)

Н.Г. Ярушкина (*jng@ulstu.ru*)

Ульяновский государственный технический университет, Ульяновск

В работе описан семантический подход резюмирования временных рядов на примере показателей состояния локальной вычислительной сети посредством использования онтологии проблемной области. Решена задача объединения различных подходов представления экспертных знаний посредством интеграции знаний продукционного характера в онтологическую модель с использованием SWRL-правил. Также в работе рассмотрена реализация данного алгоритма в программной системе анализа временных рядов.

Введение

В настоящий момент для разработки систем, основанных на знаниях, актуальной является задача объединения различных репрезентативных подходов с целью обеспечения наиболее полного представления знаний в рассматриваемой области. В рамках разработки онтологии анализа состояния локальной вычислительной сети (ЛВС), были исследованы особенности продукционного и онтологического подходов к представлению знаний соответствующей предметной области, а также был получен опыт по их совмещению в онтологической модели с помощью использования набора SWRL-правил.

Экспертные знания, используемые при решении задач логического вывода с привязкой к конкретной проблемной области, могут быть представлены в двух форматах:

1. В виде owl-онтологии, содержащей формальное описание предметной области в терминах онтологии логики принятия решений;

2. В виде набора SWRL-правил, определяющих логику взаимодействия терминов owl-онтологии, с которыми оперирует машина логического вывода.

Одно из требований к онтологиям, используемым в качестве формы хранения знаний, заключается в том, чтобы содержащиеся в них знания были «доступны» для машинной обработки, в частности, для автоматизированного логического вывода новых знаний из уже имеющихся. Для этого требуется, чтобы язык, на котором формулируются онтологии, имел точную семантику, а соответствующие логические проблемы были разрешимы (и имели практически допустимую вычислительную сложность) [Афанасьева и др., 2011].

1. Онтологическая и продукционная модели представления знаний

Формальная семантика OWL описывает, как получить логические следствия, имея такую онтологию, т.е. получить факты, которые не представлены в онтологии буквально, но следуют из ее семантики.

OWL (англ. Web Ontology Language) – язык описания онтологий, позволяющий описывать классы и отношения между ними. В основе языка – представление действительности в модели данных «объект – свойство». OWL является переформулировкой дескрипционной логики (ДЛ) с использованием синтаксиса XML [Ярушкина и др., 2013].

Поскольку существует много ДЛ, различающихся как по выразительной силе, так и по вычислительной сложности, это привело к тому, что в языке OWL имеется несколько вариантов.

OWL-DL соответствует разновидности дескрипционной логики SHOIN(D), что означает:

- S – это ALC (это набор базовых классов, имен свойств и имен объектов, является подмножеством DL) в связке с транзитивными свойствами;
- H – иерархия свойств;
- O – номиналы (так называемые классы единичных объектов);
- I – обратные свойства;
- N – числовые ограничения области значений свойств;
- (D) – типы данных, свойства, которые связывают объекты с типами данных [Митрофанова и др., 2008].

В рамках решения задачи резюмирования временных рядов (ВР), характеризующих соответствующие состояния тех или иных

b – консеквент (следствие) правила, состоящий из одного атома. Выражение в SWRL, т.е. SWRL-правило, состоит из головы и тела: голова – это консеквент продукционного правила и может состоять только из одного атома, а тело – это антецедент, который может состоять из нескольких атомов.

Процесс вывода новых фактов из содержащихся в онтологии осуществляется посредством использования резонера (рассуждателя), поддерживающего соответствующий вид логики.

Принципы работы механизма резонинга на основе SWRL-правил имеют ряд преимуществ перед остальными подобными технологиями:

1. Правила SWRL не содержат конкретных объектов, а только ссылаются на них, что дает возможность применять одно и то же правило к ряду групп объектов;

2. Правила SWRL могут быть добавлены к OWL-описанию, т.е. включены в онтологию;

3. Написание и «чтение» правил удобнее, если для этого существует специальный язык [Соловьев и др., 2006].

В рамках решаемой задачи онтологического анализа временных рядов, характеризующих состояние ЛВС, был разработан набор продукционных правил в формате SWRL, включенных в файл онтологии проблемной области.

В общем виде разработанные SWRL-правила выглядят следующим образом:

Ситуация(?x) \wedge включает Показатель(?x, Фоновая_нагрузка) \wedge имеет Тенденцию (Фоновая_нагрузка, Рост) \wedge включает Показатель(?x, Утилизация_канала_связи) \wedge имеет Тенденцию (Утилизация_канала_связи, Спад) \wedge включает Показатель (?x, Число_коллизий_в_сети) \wedge имеет Тенденцию (Число_коллизий_в_сети, Рост) \rightarrow предполагает(?x, Проблема_архитектуры)

Набор правил, которыми оперирует машина вывода информационной системы, можно разделить на две группы:

1. Группа продукционных правил, консеквентом (следствием) которых являются конкретные рекомендации по корректировке структуры ЛВС, являющиеся объектом класса «Рекомендации» онтологии предметной области.

2. Группа правил, следствием которых является логическое

присвоение одному из объектов-характеристик сети конкретного значения. Далее это значение выступает в роли входных данных других продукционных правил в процессе генерации логического вывода рекомендации.

Таким образом, в рамках обеспечения процесса семантического анализа временных рядов первоочередной является задача объединения различных репрезентативных подходов – онтологического и продукционного – с целью обеспечения наиболее полного представления знаний в рассматриваемой области.

2. Взаимодействие онтологического анализа и механизмов логического вывода

Результатом успешного решения поставленной задачи стала разработанная методика взаимодействия онтологии проблемной области и системы продукционных правил, в полной мере реализованная в программном продукте TSAalyzer 1.0.

Основная последовательность действий в процессе логического вывода рекомендации на основе взаимодействия разработанной OWL-онтологии и системы продукционных SWRL-правил включает в себя следующие этапы:

1. Извлечение данных в виде RDF-троек;
2. Проверка входных данных на наличие в онтологии соответствующих классов, отношений, объектов;
3. Запрос на выполнение набора правил анализа ситуации и вывода рекомендации;
4. Запрос на получение связанных данных (обращение к родителям объектов);
5. Запрос на выполнение правил с учетом полученных связанных данных;
6. Вывод результата анализа в виде объекта разработанной OWL-онтологии.

Непосредственная взаимосвязь самой онтологии предметной области и набора SWRL-правил в процессе логического вывода осуществляется с помощью формирования запросов к онтологии, генерируемых системой анализа при выполнении набора правил.

В качестве модуля разработанной системы TSAalyzer 1.0, обеспечивающего доступ к онтологии и возможность ее обработки, используется Java-фреймворк Apache Jena, который обеспечивает

программную среду для работы с данными в форматах RDF, RDFS, и OWL, а также поддерживающий возможность формирования запросов к онтологии на языке SPARQL.

Ядром разработанной системы, обеспечивающей логический вывод, является резонер (рассуждатель) Pellet, позволяющий обрабатывать знания в OWL-формате, основанном на дескрипционной логике. Сам процесс логического вывода, обеспечивающего выдачу пользователю экспертного заключения о состоянии тестируемой ЛВС по входным значениям характеристик загрузки сети, представленным в виде ВР, а также набору ее технических характеристик, осуществляется посредством выбора, удовлетворяющего условиям SWRL-правила, включенного в OWL-онтологию. Набор входных данных, поступающих в систему анализа ВР, представлен выражением 2.1:

$$I = \langle C, N, P, D, T \rangle, \quad (2.1)$$

где

С – множество технических характеристик рассматриваемой системы (Топология_сети, СКС, Маршрутизация и т.д.);

Н – множество характеристик работоспособности системы (Время_реакции_ПО, Загрузка_процессора_сервера и т.д.);

Р – множество свойств объектов (включает Показатель, имеет-Топологию и т.д.);

Д – множество свойств типа данных (имеет Внутреннюю Пропускную Способность, имеет Значение и т.д.);

Т – нечеткое множество характера тенденции (Спад, Рост и т.д.).

Входные данные, поступающие в систему онтологического анализа, унифицируются посредством преобразования к виду RDF-троек («субъект» – «отношение» – «объект»).

3. Анализ результатов вычислительных экспериментов на ЛВС ЦРЭМТ УлГТУ

Для проверки адекватности алгоритма интеграции онтологической и продукционной моделей представления экспертных знаний, а также корректности работы информационной системы TSAAnalyzer, реализующей данный алгоритм, был проведен ряд экспериментов, в рамках которых были смоделированы возможные проблемные ситуации, возникающие в процессе работы ЛВС при искусственном повышении загрузки каналов связи.

В рамках эксперимента было смоделировано 5 проблемных ситуаций снижения производительности ЛВС, причины появления которых необходимо было распознать, используя основанную на предложенной методике взаимодействия онтологического анализа и логических механизмов информационную систему TSA analyzer и разработанную экспертом OWL-онтологию с включением набора SWRL-правил.

Архитектура ЛВС, моделируемые проблемы в процессе искусственного повышения трафика в сети и семантический вывод системы TSA analyzer, содержащий рекомендацию по решению проблем, представлены в таблице 1.

Таблица 1

Характеристики ЛВС	Моделируемые проблемы	Вывод системы TSA analyzer
<p>Топология – «звезда», количество рабочих станций – 14, ОС сервера – Windows Server 2008 R2, сетевой коммутатор – D-Link DGS-3420-28SC, 20 портов, объем оперативной памяти коммутатора – 2 Mb, режим работы – 100 Мб/с Full Duplex, режим борьбы с коллизиями – CSMA/CD, QoS вкл, тип коммутационных шнуров – UTP, категория – 5Е, разъемы – RJ-45.</p>	<p>Быстрый рост фоновой нагрузки, рост времени реакции пользовательского ПО</p>	<p>Причина – перегруженность канала связи. Требуется изменение архитектуры сети. Число станций в перегруженных доменах следует уменьшить, а станции, создающие наибольшую загрузку, – подключить к выделенным портам коммутатора</p>
	<p>Активация режима Half Duplex, рост фоновой нагрузки</p>	<p>Причина – несоответствие характеристик сетевого коммутатора объему трафика. Требуется замена коммутатора или изменение его настроек</p>
	<p>Повреждение одного из коммутационных шнуров, рост числа коллизий</p>	<p>Причина – неполадки с подключением коммутационных шнуров. Причина – неправильная организация заземления компьютеров, включенных в локальную сеть</p>

Продолжение таблицы 1

	Наличие дефектов у портов коммутатора, рост числа ошибок CRC, рост фоновой нагрузки	Причина – неполадки работы портов сетевого коммутатора или повреждение коммутационного шнура, ведущего к проблемному хосту
	Наличие дефектов сетевой платы, быстрый рост числа локальных и удаленных коллизий	Причина – проблема в настройках сетевой платы сервера домена, проверьте целостность сетевой карты и правильность ее настройки

В четырех случаях из пяти система наиболее близко сформулировала проблемное место рассматриваемой ЛВС, в одном из экспериментов (в третьем) рекомендация включила в себя несколько вариантов исправления проблемы ввиду недостаточной полноты поступающих в систему данных состояния тестируемой ЛВС.

Помимо этого, в трех случаях из пяти рекомендации по корректировке архитектуры сети генерировались при последовательной активизации двух и более SWRL-правил. Данный подход моделирует процесс принятия решения человеком.

Стоит отметить, что дальнейшая детализация правил вывода с целью получения более точного и адекватного значения рекомендации может быть реализована по двум сценариям:

1. За счет увеличения количества атомов в antecedente правил. Таким образом, каждое правило будет охватывать большее число параметров, от которых может зависеть производительность сети. Главным недостатком такого пути является заметное замедление процесса обработки модифицированных правил.

2. За счет разбиения правил на более короткие с увеличением количества SWRL-конструкций, следствием которых будет присвоение конкретного значения отдельному параметру сети с его последующим участием в процессе логического вывода в качестве входных данных другого правила. Данный подход является рациональным с точки зрения временных затрат и вариативности исходов процесса вывода.

Заключение

Таким образом, предлагаемый в работе семантический подход резюмирования ВР на основе онтологии проблемной области является не только наиболее удобным с точки зрения восприятия процесса и результатов анализа человеком, но и достаточно универсальным, исходя из возможности смены позиции рассмотрения поставленной задачи при помощи замены онтологии необходимой предметной области.

Данная методика объединяет в себе два подхода хранения и обработки знаний – онтологический и продукционный – и позволяет формировать лексическое описание результатов анализа временных рядов с привязкой к конкретной предметной области, обновляя структуру онтологии за счет введения и выполнения новых продукционных правил.

СПИСОК ЛИТЕРАТУРЫ

[Афанасьева и др., 2011] Афанасьева Т.В., Ярушкина Н.Г. Нечеткий динамический процесс с нечеткими тенденциями в анализе временных рядов//Вестник Ростовского государственного университета путей сообщения. – 2011. – Т. 3. – С. 7–16.

[Митрофанова и др., 2008] Митрофанова О.А., Константинова Н.С. Онтологии как системы хранения знаний / Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы», 2008. – 54 с.

[Соловьев и др., 2006] Соловьев В.Д., Добров Б.В., Иванов В.В., Лукашевич Н.В. Онтологии и тезаурусы. Учебное пособие. – Казань, Москва.– 2006.

[Ярушкина и др., 2013] Ярушкина Н.Г., Мошкин В.С. Онтологический подход к анализу временных рядов//Интегрированные модели и мягкие вычисления в искусственном интеллекте. – 2013. – №2. – С.529–537.

УДК 004.891.2

**ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ПОДДЕРЖКИ
ПРИНЯТИЯ РЕШЕНИЙ ПРИ ОБОСНОВАНИИ ВЫБОРА
КОНСТРУКЦИОННЫХ МАТЕРИАЛОВ¹**

А.Ф. Берман (*berman@icc.ru*)

Г.С. Малтугуева (*gama@icc.ru*)

А.Ю. Юрин (*iskander@icc.ru*)

Институт динамики систем и теории управления СО РАН, Иркутск

В работе описываются интеллектуальная система поддержки принятия решений (ИС ППР) при обосновании выбора конструкционных материалов, а также алгоритмическое обеспечение, основанное на совместном применении прецедентного подхода и методов многокритериального выбора. Рассмотрена модель прецедента, приведен пример применения ИС ППР.

Введение

Для решения широкого круга инженерных задач, связанных с многокритериальной оптимизацией параметров, многовариантным планированием и последующим выбором, применяются различные методы многокритериального выбора (оптимизации) [Гаибова, 2004; Дилигенский и др., 2004; Ройзензон, 2005; Соболев и др., 2006; Денисова, 2013а]. В результате проведенного аналитического обзора было отмечено, что при решении конструкторских задач

¹ Работа выполнена при частичной финансовой поддержке РФФИ (проект №14-01-31321).

в области машиностроения, например, при проектировании нового изделия (детали), требуется решить задачу выбора конструкционного материала. Данная задача является многовариантной, так как для изготовления какой-либо детали можно выбрать некоторое множество материалов, удовлетворяющих условиям ее эксплуатации, изготовления и обслуживания в составе конструкции или машины. Как правило, при этом необходимо учитывать целый ряд условий: назначение детали; условия эксплуатации, технологию изготовления и ремонта, стоимость и доступность и др., которые неоднозначно оцениваются специалистами, принимающими решения.

В настоящее время задача выбора конструкционных материалов решается либо «по аналогии» [Машиностроение, 2004], либо по описанию эксплуатационных и технологических свойств материала [Справочник, 2005], либо методом экспертного оценивания [Машиностроение, 2004] с дальнейшим применением методов многокритериальной оптимизации. При этом отсутствует программное обеспечение, автоматизирующее решение данной задачи.

Целью работы является создание ИС ППР для выбора конструкционного материала. В качестве методологической основы процесса принятия решений предлагается использовать объединение подхода «по прецедентам» при принятии решений (в качестве метода поиска возможных вариантов) и методов многокритериального выбора (для получения упорядоченного перечня вариантов).

1. Постановка задачи и методы решения

Принятие решений всегда сопряжено с необходимостью оценивания всех возможных (доступных) вариантов по набору критериев, каждый из которых измеряется по собственной шкале.

Для решения задачи выбора конструкционных материалов необходимо упорядочить все пригодные конструкционные материалы по убыванию степени их предпочтительности, которая основывается на значениях их характеристик и максимальной приближенности к заданным условиям и целям (требования к жаропрочности, коррозионной стойкости, трещиностойкости, наличие в требуемом количестве и по приемлемой цене у ближайшего производителя).

Обозначим через $A=(A_1, A_2, \dots, A_n)$ множество всех возможных конструкционных материалов, из которых необходимо выбрать

k вариантов $X=(X_1, X_2, \dots, X_k)$, $X \in A$, удовлетворяющих заданной цели и имеющимся условиям (ограничениям). После этого необходимо упорядочить отобранные варианты в порядке убывания степени их предпочтительности для лица, принимающего решение (ЛПР), которая основывается на оценках по набору критериев $C=(C_1, C_2, \dots, C_m)$. В виде критериев можно представить требуемые характеристики.

Для решения поставленной задачи предлагается применить: рассуждения на основе прецедентов (case-based reasoning) [Aamodt et al., 1994], что обеспечит поиск и извлечение материалов, подходящих под определенные условия функционирования (эксплуатации); методы многокритериального выбора [Петровский, 2009] для упорядочения выбранных конструкционных материалов в соответствии со значениями их характеристик и мерой близости аналогов.

1.1. Прецедентный подход

В соответствии с постановкой задачи предлагается хранить информацию о конструкционных материалах (и опыте их применения) в библиотеке (базе) прецедентов. Под прецедентом понимается компактное описание знаний о конструкционном материале, содержащее наиболее важные параметры и свойства рассматриваемого объекта. Как правило, прецедент состоит из описания проблемной ситуации и совокупности действий, предпринимаемых в данной проблемной ситуации (ее решения).

Ниже приведен фрагмент описания прецедента:

Описание	
материал:	20Х3МВФ
жаропрочность:	да
температура применения:	до 500–560 °С
срок работы:	от 1000 до 10000 ч.
коррозионная стойкость:	незначительная
трещиностойкость:	да
технологичность:	низкая
стоимость:	высокая

Решение		
	конструктивный элемент:	труба в трубе
условия эксплуатации:	давление:	220 МПа
	температура:	300 °С
	среда:	слабоагрессивная

Все прецеденты в базе (библиотеке) прецедентов индексируются [Николайчук и др., 2009]. Поиск осуществляется путем вычисления расстояния между индексами при помощи метрики Миньковского [Aamodt et al., 1994].

В качестве входных параметров для процедуры поиска используются: вид детали, условия ее функционирования (эксплуатации).

В качестве результата процедуры извлечения пользователь получает набор аналогов, упорядоченных в соответствии с мерой подобия (близости) описаний. Каждый аналог содержит описание определенного материала.

Для дальнейшей обработки извлеченных вариантов применяются методы многокритериального выбора.

1.2. Методы многокритериального выбора

Методы многокритериального выбора предназначены для обработки критериальных оценок альтернатив и результат их работы может иметь один из трех видов: единственная (лучшая) альтернатива, упорядочение всех альтернатив и классификация на группы. В нашем случае требуется упорядочить имеющийся перечень вариантов, поэтому в дальнейшем будут рассматриваться только методы упорядочения (ранжирования) всех альтернатив.

Входная информация для всех методов многокритериального выбора представляет собой совокупность оценок всех допустимых альтернатив по некоторым критериям, каждый из которых измеряется по некоторой шкале значений (качественная, количественная).

В рассматриваемой задаче предлагается в качестве критериев рассматривать характеристики, параметры конструкционных материалов, хранящиеся в «Описании» прецедентов. Дополнительным критерием будет выступать оценка близости прецедента.

После оценивания всех альтернатив по каждому критерию осуществляется обработка полученных критериальных оценок, в

процессе которой от лица, принимающего решение, может потребоваться дополнительная информация. Существует множество методов многокритериального выбора, позволяющих обрабатывать разнородную информацию, например, метод анализа иерархий, методы вербального анализа, метод уступок, методы свертки, генетические алгоритмы, метод исследования пространства признаков и другие [Саати, 1993; Земских и др., 2004; Васильев и др., 2005; Соболев и др., 2006; Подиновский, 2007; Петровский, 2009; Денисова, 2013b].

Рассмотрев примеры успешного применения многометодного подхода [Белецкая, 2005; Ройзензон, 2008], было решено применить его и в рассматриваемой задаче.

На рис. 1 представлена принципиальная схема работы совместного использования прецедентного подхода и методов многокритериального выбора.

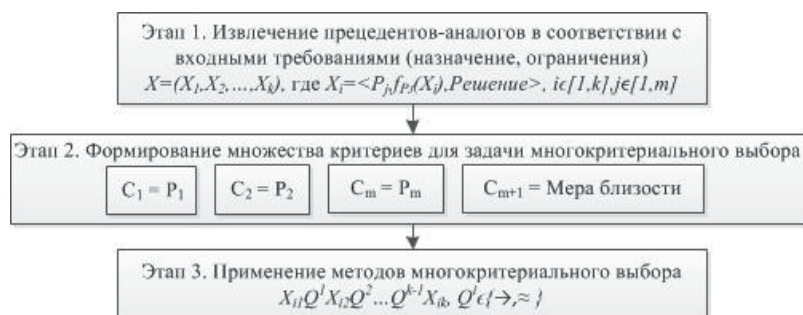


Рис. 1. Принципиальная схема предлагаемого подхода

Окончательное решение, предъявляемое лицу, принимающему решение, представляет собой список вариантов, упорядоченных по убыванию степени их предпочтительности, и является результатом согласования оценок альтернатив по всем критериям и с учетом степени близости выбранных материалов к текущей ситуации.

2. Интеллектуальная система поддержки принятия решений

Процесс поддержки принятия решений может быть автоматизирован при помощи компьютерной системы поддержки принятия решений (СППР).

Основные функции СППР:

- управление библиотекой (базой) прецедентов, включая добавление, изменение, удаление описания материалов;
- поиск и извлечение прецедентов по заданному описанию;
- обработка результатов поиска при помощи методов многокритериального выбора;
- объяснение и отображение результатов.

С целью повышения эффективности процесса согласования критериальных оценок альтернатив в состав СППР могут быть введены программные модули, реализующие другие методы многокритериального выбора [Петровский, 2009].

Модули реализованы в виде отдельных программных компонентов – динамических библиотек (DLL), что позволяет их использовать в составе других программных систем.

Архитектура СППР включает:

- модуль управления библиотеками прецедентов;
- модуль поиска решения по прецедентам;
- модули многокритериального выбора (МАИ, ЗАПРОС I, ЗАПРОС II, АРАМИС);
- подсистему отображения результатов;
- подсистему объяснения полученных результатов.

3. Пример применения

Рассмотрим пример применения ИС ППР для выбора материала для изготовления труб, предназначенных для работы в качестве теплообменных под высоким внутренним давлением, при температуре 300 °С, под воздействием на наружную поверхность слабо агрессивной среды, при требовании значительного ресурса и трещиностойкости в составе реактора типа «труба в трубе».

База прецедентов содержит описание материалов [Кузнецова и др., 1999; Справочник, 2005]. Запрос к базе прецедентов включает следующие данные:

Конструктивный элемент: труба в трубе

Условия эксплуатации:

Давление: 220 МПа

Температура: 300 °С

Среда: слабоагрессивная

Требования: жаропрочность (да), трещиностойкость (да), технологичность (высокая).

В результате выполнения запроса был получен следующий перечень аналогов (возможных материалов):

Материал	20Х3МВФ	25Х2М1Ф	30ХМА	18Х3МВ	12ХГ
Обозначение	X_1	X_2	X_3	X_4	X_5
Близость	0,85	0,85	0,83	0,83	0,73

Каждый элемент из множества отобранных материалов X характеризуется следующими оценками по критериям (пс – полностью соответствует, чс – частично соответствует, нс – не соответствует, н – низкая, с – средняя, в – высокая):

Критерии		Альтернативы				
		X_1	X_2	X_3	X_4	X_5
Давление	C_1	пс	пс	пс	пс	чс
Температура	C_2	пс	пс	пс	пс	пс
Среда	C_3	пс	пс	пс	пс	пс
Теплостойкость	C_4	пс	пс	чс	пс	чс
Жаростойкость	C_5	пс	пс	пс	пс	чс
Жаропрочность	C_6	пс	пс	чс	пс	чс
Трещиностойкость	C_7	пс	пс	пс	пс	чс
Технологичность	C_8	н	н	в	н	в
Стоимость	C_9	в	в	с	с	н
Близость	C_{10}	0,85	0,85	0,83	0,83	0,73

Для упорядочения альтернатив (материалов) по убыванию степени соответствия всей совокупности их критериальных оценок поставленной задаче применимы методы многокритериального выбора. В результате применения методов обработки равнозначных критериев (без назначения весов, уровня значимости) будет получено упорядочение следующего вида: $X_1 \approx X_2 \approx X_4 \rightarrow X_3 \rightarrow X_5$. Полученное упорядочение уточняет результат, полученный на предыдущем этапе (извлечение прецедентов-аналогов), и является подтверждением достоверности работы прецедентного подхода.

Для построения строгого упорядочения на множестве альтернатив необходимо либо назначать веса критериев, либо устанавливать уровень их значимости. Кроме того, в ряде методов многокри-

териального выбора возможно обобщение критериальных оценок нескольких экспертов [Петровский, 2009].

Заключение

Рассмотрен способ повышения эффективности и качества принятия инженерных решений, в том числе при обосновании выбора конструкционных материалов. Способ основан на применении методов поиска решений на основе прецедентов и теории многокритериального выбора.

Необходимо отметить ограничения реализованного в ИС ППР способа: прецедентный метод применим только при наличии представительной базы прецедентов, содержащей опыт применения конструкционных материалов. Кроме того, применение других метрик при поиске аналогов может приводить к иным результатам на этапе извлечения.

В свою очередь, применение методов многокритериального выбора позволяет уточнить результат, полученный на этапе извлечение прецедентов-аналогов, и подтвердить его достоверность.

СПИСОК ЛИТЕРАТУРЫ

[Aamodt et al., 1994] Aamodt A., Plaza E. Case-based reasoning: Foundational issues, methodological variations, and system approaches // *AI Communications*, vol. 7 (1994), no. 1.

[Белецкая, 2005] Белецкая С.Ю. Оптимизация принятия решений в САПР на основе интеграции многовариантного моделирования и адаптивной мультикомпонентной поисковой среды // Автореферат диссертации на соискание ученой степени доктора технических наук: 05.13.12. – Воронеж, 2005.

[Васильев и др., 2005] Васильев С.Н., Котлов Ю.В. Технология поиска компромисса при индивидуальном и групповом выборе // Труды Всероссийской конференции «Равновесные модели экономики и энергетики». Иркутск – Северобайкальск, 3–7 июля 2005 г., т. Математическая экономика. – Иркутск: ИСЭМ СО РАН, 2005.

[Гайбова, 2004] Гайбова Т.В. Многокритериальная оптимизация инвестиционных проектов развития промышленных предприятий // Автореферат на соискание ученой степени кандидата технических наук: 05.13.01. – Самара: Самарский государственный технический университет, 2004.

[Денисова, 2013а] Денисова Л.А. Моделирование и оптимизация системы регулирования питания парогенератора энергоблока АЭС // *Автоматизация в промышленности*, 2013. – №7.

[Денисова, 2013b] Денисова Л.А. Автоматизация параметрического синтеза системы регулирования на основе многокритериальной оптимизации с использованием генетического алгоритма // Автоматизация в промышленности, 2013. – №12.

[Дилигенский и др., 2004] Дилигенский Н.В., Дымова Л.Г., Севастьянов П.В. Нечеткое моделирование и многокритериальная оптимизация производственных систем в условиях неопределенности: технология, экономика, экология. М.: «Издательство Машиностроение», 2004.

[Земских и др., 2004] Земских Л.В., Самаров Е.К., Жданов А.А., Бабкова В.В. Применение генетических алгоритмов для оптимизации адаптивной системы управления мобильного робота на параллельном вычислительном комплексе // Труды Института системного программирования РАН. – М.: ИСА РАН, 2004. – Том 7.

[Кузнецов и др., 1999] Сосуды и трубопроводы высокого давления: Справочник под редакцией А.М. Кузнецова и В.И. Лившица. Изд. 2-е. Иркутск. Типография №1, 1999.

[Машиностроение, 2004] Машиностроение. Энциклопедия / Ред. совет: К.В. Фролов (пред.) и др. – М.: Машиностроение. Машины и аппараты химических производств. Т. IV-12 / М.Б. Генералов и др. 2004.

[Николайчук и др., 2009] Николайчук О.А., Юрин А.Ю. Применение прецедентного подхода для автоматизированной идентификации технического состояния деталей механических систем // Автоматизация и современные технологии, 2009. – №5.

[Петровский, 2009] Петровский А.Б. Теория принятия решений. – М.: Академия, 2009.

[Подиновский, 2007] Подиновский В.В., Ногин В.Д. Парето-оптимальные решения многокритериальных задач. 2-е изд. – М.: Физматлит, 2007.

[Ройзензон, 2005] Ройзензон Г.В. Многокритериальный выбор вычислительных кластеров // Труды Института системного программирования РАН. – М.: ИСА РАН, 2005. – Том 12.

[Ройзензон, 2008] Ройзензон Г.В. Интерактивные методы снижения размерности признакового пространства в задачах многокритериального принятия решений // Автореферат диссертации на соискание ученой степени кандидата технических наук: 05.13.10. – М.: ИСА РАН, 2008.

[Саати, 1993] Саати Т. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993.

[Соболь и др., 2006] Соболь И.М., Статников Р.Б. Выбор оптимальных параметров в задачах со многими критериями. – 2-е изд. – М.: Дрофа, 2006.

[Справочник, 2005] Справочник по конструкционным материалам: Справочник / Б.Н. Арзамасов и др.; под ред. Б.Н. Арзамасова. – М.: МГТУ им. Н.Э. Баумана, 2005.

УДК 004.89:004.424.4:519.87

ПРОТОТИП ПРОГРАММНОЙ РЕАЛИЗАЦИИ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ ДОКУМЕНТАЛЬНОГО ПОИСКА¹

В.К. Иванов (*mtivk@tstu.tver.ru*)

Б.В. Палюх (*pboris@tstu.tver.ru*)

П.И. Мескин (*pavel.meskin@gmail.com*)

Тверской государственный технический университет, Тверь

В статье представлен прототип программной реализации генетического алгоритма, который может использоваться для формирования эффективных поисковых запросов и отбора релевантных результатов при выполнении документального тематического поиска. Рассматриваются целевая функция, основные шаги и параметры алгоритма. Описываются компоненты программного обеспечения. Приведены некоторые результаты предварительных исследований алгоритма.

Введение

Каждый пользователь Интернета с помощью поисковых систем искал в сети какие-либо конкретные факты и, как правило, находил либо точно то, что искал, либо что-то близкое. Результат такого поиска – описание объекта, события или явления с заданными значениями их свойств.

В других случаях мы ищем целые категории скоординированной информации в некотором тематическом сегменте, а не отдельные информационные объекты с заданными характеристиками (представителей этих категорий). Результатом, кроме набора фактов, следует считать сведения о ретроспективе, перспективе, взаимосвязях найденных информационных объектов, о текущих и вероятных трендах.

При выполнении тематического поиска неизбежно возникает ряд вопросов. Как совместно оценить релевантность документов, найденных разными запросами? Является ли ранжирование результатов корректным с позиций ожиданий пользователя? Все ли

¹ Работа выполнена при финансовой поддержке РФФИ (проект № НК13-07-00342).

результаты, соответствующие ожиданиям пользователя, попали в число доступных для непосредственной оценки? Как отфильтровать документы, не относящиеся по сути к искомой тематике? Могут ли быть обнаружены эффективные решения, которые относятся к другим областям применения, но могут быть успешно использованы в данной области? Невозможно однозначно ответить на эти вопросы в рамках тривиальных решений.

Для получения эффективного множества поисковых запросов мы предлагаем использовать подход, основанный на генетических алгоритмах, область применения которых может быть достаточно широкой (см., например, системное изложение в [Курейчик и др., 2006]). Подход предусматривает организацию эволюционного процесса, а) формирующего устойчивую и эффективную популяцию поисковых запросов, б) образующего соответствующий поисковый образ документов или семантического ядра, в) приводящего к получению релевантного искомого множества документов.

Примененные подходы и технологии

Представленная в данной статье разработка, являясь частью проекта «Интеллектуальная распределенная система информационной поддержки инноваций в науке и образовании» [Палюх и др., 2013], касается решения задач фильтрации данных [Hanani et al., 2001], [Маннинг и др., 2011] на основе непосредственного использования контента. Для нас представляли существенный интерес исследования видов поиска в Web [Broder, 2002] и различных вариантов альтернативных техник поиска, таких как использование паттернов [Zhu et al., 2013], с возможностью специфицировать информационные потребности в виде набора релевантных документов (а не набора ключевых слов) и оптимальное совместное использование нескольких информационных источников для получения набора запросов [Wu et al., 2011] и [Bendersky et al., 2012].

Отметим, что для различных программных платформ разработано довольно большое количество реализаций генетических алгоритмов. Есть как не потерявшие своей актуальности разработки [Wall, 1996], так и современные реализации, например, GeneHunter (<http://www.wardsystems.com>) или Genetic Algorithm Framework for Net (<http://johnnewcombe.net/gaf>). Некоторые авторы делают попытки унификации и стандартизации подходов к разработкам

[Сергиенко и др., 2010], а другие ориентируются на специальное применение генетических алгоритмов [Подлазова, 2008].

При разработке обсуждаемого в статье алгоритма было принято решение об оригинальной реализации основных генетических операций. Причины следующие:

1. Особенности вычисления значений целевой функции.
2. Представление поисковых запросов как хромосом, гены которых (ключевые слова) выражены значениями номинальной шкалы.
3. Специфические интерпретации базовых генетических операций – обмен понятийными элементами и использование синонимии.
4. Необходимость исследований, анализа и выполнения потенциальных улучшающих модификаций алгоритма.
5. Планируемое использование генетического алгоритма в интеллектуальном поисковом приложении для мобильных устройств.

Целевая функция

Для каждого i -го результата запроса целевая функция (пригодность) может быть определена как $w_i = f(p, r, s)$, где

p – средний номер позиции адреса документа в списке первых P результатов выполненных поисковых запросов (учитываются только те списки результатов, где данный адрес присутствует);

r – количество появлений адреса документа в результатах выполнения N поисковых запросов. Отметим, что $r \leq N$ и $r = N$, если адрес документа появился в результатах выполнения всех запросов;

s – семантическая близость текста найденного документа (по крайней мере заголовка и сниппета) и поискового образа документов (исходного множества ключевых слов). В качестве альтернативы используются эталонные тексты, адаптивно формируемые в ходе выполнения алгоритма. Для вычисления s использована модификация модели векторного пространства документов, в которой каждый документ интерпретируется как вектор $\vec{v}(d) = (w_{1,d}, w_{2,d}, \dots, w_{M,d})$, где $w_{i,d} = tf_{i,d} * idf_{i,d}$. Здесь $tf_{i,d}$ – частота использования термина в документе, $idf_{i,d} = \log(P + 1/P_i)$, где P – общее число документов в SERP), P_i – число документов, содержащих данный термин, M –

число терминов в P документах. Близость текстов s интерпретируется как косинусная мера близости $s(d_1, d_2)$.

Значения p, r, s нормированы на диапазон от 0 до 1. Ранг каждого результата запроса вычисляется как средневзвешенное значение $w_i = f_5 * p + f_6 * r + f_7 * s$, где f_5, f_6, f_7 – весовые коэффициенты, являющиеся параметрами алгоритма, $f_5 + f_6 + f_7 = 1$.

Значение целевой функции для запроса \bar{w} – средний ранг результатов этого запроса, а значение целевой функции для популяции запросов \bar{W} – средний ранг запросов этой популяции.

Описание алгоритма

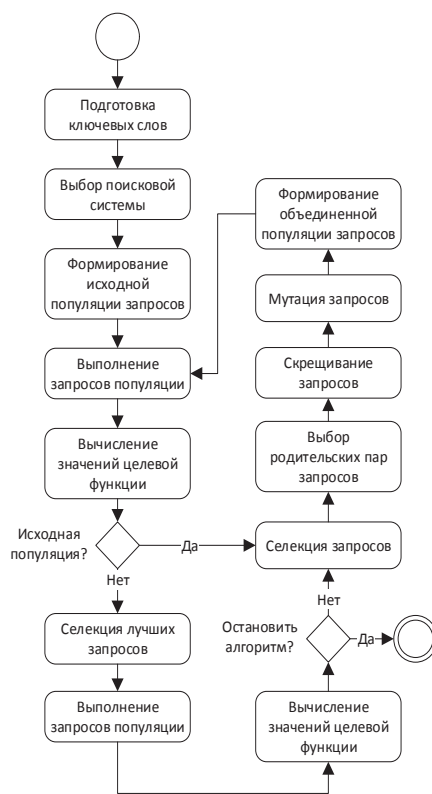


Рис. 1. Генетический алгоритм формирования эффективного множества запросов

На рис. 1 приведена схема генетического алгоритма формирования эффективного множества запросов. Ниже кратко описаны основные шаги алгоритма; более подробное описание см. в [Иванов, 2013].

1. Подготовка ключевых слов. Множество ключевых слов $K = \{k_j\}$, формирует поисковый образ множества документов заданной тематики.

2. Выбор поисковой системы. Используется любая поисковая система (или несколько одновременно), имеющая API: Bing, Google и др.

3. Формирование исходной популяции запросов. Каждый поисковый запрос есть совокупность ключевых слов. Исходная популяция из N запросов – это множество $Q = \{q_i\}$, $|Q| = N$, где $q_i = \{k_1, k_2, \dots, k_m\}$ – случайно образованная комбинация ключевых понятий из K , $k_j \in K$ и для любой пары (k_r, k_s) из K справедливо $k_r \neq k_s$.

4. Выполнение запросов популяции. Результатом являются возможно пересекающиеся множества дескрипторов найденных документов (заголовки, описание, адрес текста).

5. Вычисление целевой функции.

6. Селекция лучших запросов. Следующее поколение формируется теми запросами, у которых $\bar{w} \geq \bar{W}$, где порог \bar{W} может быть средним значением или медианой по текущей популяции.

7. Выбор родительских пар запросов. Используется генотипный аутбридинг. Расстояние между родителями вычисляется как $\Delta\bar{w} = \bar{w}_1 - \bar{w}_2$.

8. Скрещивание запросов. Реализуется одноточечным кроссинговером – обменом ключевыми словами или их группами между запросами. Особенность реализации – ключевое слово k_i запроса-родителя замещается синонимом слова из запроса-родителя $k_{st} \in S_i$, где S_i – множество синонимов слова k_i (генерируется существенно больше потомков при сохранении семантики родителей).

9. Мутация запросов. Наиболее адекватная операция мутации в нашем подходе – замена случайно выбранного слова запроса на его синоним. Невозможно применение присоединения или вставки нового термина, удаления термина, обмена местами терминов.

10. Формирование объединенной популяции. Используется элитарный отбор. Лучшие решения не теряются.

11. Остановка алгоритма. Общее условие остановки – стабильность популяции. Например, достижение среднееквadraticным отклонением значений целевой функции пороговой величины. Есть варианты использования заданного или предельного числа проходов алгоритма.

Параметры алгоритма

Основные параметры: g_1 – поисковая система (например, Bing или Google); g_2 – количество запросов в генерируемых популяциях; g_3 – количество ключевых слов в каждом генерируемом запросе; g_4 – исходный набор ключевых слов и понятий.

Параметры целевой функции: f_1 – количество результатов поиска, возвращаемых запросом; f_2 – количество результатов поиска, возвращаемых после ранжирования результатов популяции; f_3 – количество результатов поиска, возвращаемых после ранжирования результатов всех популяций; f_4 – коэффициент, учитывающий рас-

положение документов на одном сервере; f_5, f_6, f_7 – веса аргументов p, r и s соответственно при расчете ранга результата поиска; f_8 – способ вычисления целевой функции для групп результатов.

Параметры скрещивания и мутации: c_1 – множитель, используемый для вычисления критерия отбора запросов-родителей: $\bar{w} > c_1 * \bar{W}$; m_1 – вероятность мутации запроса.

Параметры завершения алгоритма: e_1 – заданное число проходов алгоритма; e_2 – предел среднеквадратичного отклонения целевой функции $\sigma < e_2$; e_3 – предельное число проходов алгоритма.

Реализация алгоритма

Основные функциональные компоненты приложения на платформе .NET Framework, реализующего алгоритм, следующие:

1. Пользовательский интерфейс на базе Windows Form.

2. Основная библиотека алгоритма GAF. Объектная модель реализует следующие основные классы: *GAF* – свойства, операции, входные и выходные данные алгоритма; *Options* – параметры алгоритма; *Population* – свойства и операции популяции; *Individual* – свойства и операции запроса популяции. Класс *Individual* имеет клоны *SelectedParents* (запросы-родители), *SelectedParentsPairs* (пары запросов-родителей), *Children* (запросы-потомки), *MutedChildren* (мутировавшие запросы); *SearchResult* – свойства и операции элемента результатов запроса; *TextSimilarity* – определяет параметры вычисления схожести текстов.

3. Модуль морфологического анализа и лемматизации запросов. Используется технология phpMorphy с бинарными словарями (<http://phpmorphy.sourceforge.net>).

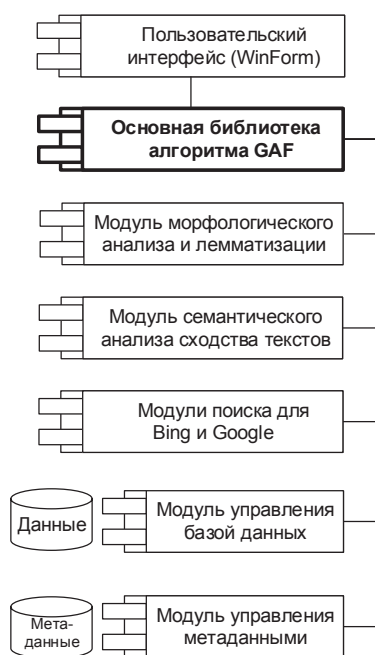


Рис. 2. Основные компоненты приложения, реализующего генетический алгоритм

4. Модуль семантического анализа сходства текстов. Используется экспериментальная платформа, описанная в [Иванов, 2012].

5. Модули поиска, обеспечивающие унификацию коллекций результатов на базе Bing Search API и Google Custom Search API.

6. Модуль управления базой данных. Обеспечивает доступ к используемым словарям и сохранение результатов работы алгоритма.

7. Модуль управления метаданными (разработан на базе XML-схем). Обеспечивает обработку параметров.

Предварительное исследование алгоритма

В экспериментах использовалась программная реализация алгоритма, описанная выше, и следующий исходный набор значений параметров: $g_1 = \text{Bing}$, $g_2 = 15$, $g_3 = 3$, $f_1 = 10$, $f_2 = 50$, $f_3 = 50$, $f_4 = 0,75$, $f_5 = 0,33$, $f_6 = 0,33$, $f_7 = 0,34$, $f_8 = \text{Average}$, $c_1 = 0$, $m_1 = 1$, $e_1 = 10$, $e_2 = 0$, $e_3 = 50$. В исходном наборе g_4 использовались термины с наибольшими весами $tf_{i,d} * idf_{i,d}$ из корпуса документов – студенческих рефератов.

На рис. 3 показаны графики зависимости \bar{W} от номера популяции при $g_4 = 50$. Видны локальные максимумы \bar{W} и точки начала относительной стабилизации \bar{W} (5–6-я популяция).

На рис. 4 показаны графики зависимости \bar{W} от g_3 (показаны числами около графиков). Видно, что увеличение g_3 в целом приводит к улучшению качества популяций.

На рис. 5 показано влияние аргументов целевой функции на ее значение. Наибольшее влияние оказывает p , наименьшее – r .

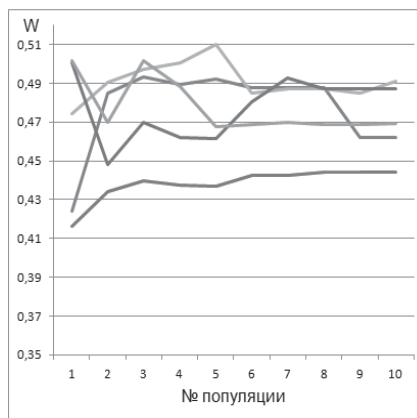


Рис. 3. Изменение \bar{W} ($g_4=50$, $e_1=10$)

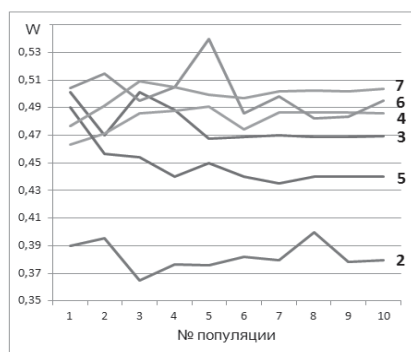


Рис. 4. Изменение \bar{W} ($g_4=50, g_3=2\dots7$)

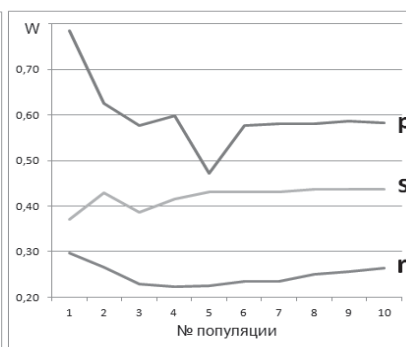


Рис. 5. Влияние аргументов на \bar{W}

Заключение

Генетический алгоритм, рассмотренный в настоящей статье, является одним из элементов программного обеспечения разрабатываемой интеллектуальной системы информационной поддержки инноваций в науке и образовании. В настоящее время проводятся исследования эффективности алгоритма с точки зрения пользователей.

Описанные подходы к организации тематического поиска могут быть успешно применены в ряде областей. Например, обзоры источников научно-технической, коммерческой и социальной информации, поиск коммерчески ценных данных, сбор информации о клиентах, определение новых областей при бизнес-планировании, конкурентный анализ и разведка, поиск инновационных решений, подбор и экспертиза учебно-методических материалов, анализ конкурсной документации и условий экспертизы проектов, подборка материалов для патентных исследований.

СПИСОК ЛИТЕРАТУРЫ

[Иванов, 2012] Иванов В.К. Критерии интегральной оценки электронных документов в системах подготовки принятия решений: статья // Вестник ТвГТУ: Вып. 22. – С. 20–26. – Тверь, 2012.

[Иванов, 2013] Иванов В.К. Основные шаги генетического алгоритма фильтрации результатов тематического поиска документов : статья // Инновации в науке. – Новосибирск, 2013. – № 25. – С. 8–15.

[Курейчик и др., 2006] Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы / Под ред. В.М. Курейчика. – 2-е изд., испр. и доп. – М.: Физматлит, 2006. – 320 с.

[Маннинг и др., 2011] Маннинг, Кристофер Д., Рагхаван, Прабхакар, Щютце, Хайнрих. Введение в информационный поиск. : Пер. с англ. – М.: ООО «И.Д. Вильямс», 2011. – 528 с.

[Палюх и др., 2013] Палюх Б.В., Иванов В.К., Сотников А.Н. Архитектура интеллектуальной системы информационной поддержки инноваций в науке и образовании // Программные продукты и системы. – 2013. – № 4. – С. 197–202.

[Подлазова, 2008] Подлазова А.В. Генетические алгоритмы на примерах решения задач раскрытия // Проблемы управления. – 2008. – №2. – С. 57–63.

[Сергиенко и др., 2010] Генетический алгоритм. Стандарт: Ч. 1: Описание стандартного генетического алгоритма (сГА) / А. Б. Сергиенко, П. В. Галушин, В. В. Бухтояров [и др.]; Сибирский гос. аэрокосмический ун-т им. М. Ф. Решетнева. – Красноярск, 2010. – 384 с.

[Bendersky et al., 2012] Bendersky M., Metzler D., Croft W. B. Effective query formulation with multiple information sources. WSDM'12 Proceedings of the fifth ACM international conference on Web search and data mining, pp. 443-452, DOI: 10.1145/2124295.2124349.

[Broder, 2002] Broder A. A taxonomy of web search, ACM SIGIR Forum Vol. 36, Issue 2, Fall 2002, pp. 3–10, DOI: 10.1145/792550.792552.

[Hanani et al., 2001] Hanani U., Shapira B., Shoval P. Information Filtering: Overview of Issues. Research and Systems, User Modeling and User-Adapted Interaction II: pp. 203–259. 2001.

[Wall, 1996] Wall M. Galib: A C++ library of genetic algorithm components. Mechanical Engineering Department, Massachusetts Institute of Technology, 87, 1996.

[Wu et al., 2011] Wu J., Ilyas I., Weddell G. A Study of Ontology-based Query Expansion, Cheriton School of Computer Science, University of Waterloo, Technical Report CS-2011-04, February 09, 2011, p. 38.

[Zhu et al., 2013] Zhu M., Xu C., Wu Y.-F.B. IFME: information filtering by multiple examples with under-sampling in a digital library environment. JCDL'13 Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries, pp. 107-110, DOI: 10.1145/2467696.2467736.

УДК 004.82:004.89:519.816

СОЗДАНИЕ ПОРТАЛА ЗНАНИЙ ДЛЯ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ ПОДДЕРЖКИ РАЗРАБОТЧИКОВ СППР¹

Г.Б. Загорулько (*gal@iis.nsk.su*)

Институт систем информатики им. А.П. Ершова
СО РАН, Новосибирск

В.Г. Итыгилов (*iti.vadim@gmail.com*)

Новосибирский государственный университет, Новосибирск

В докладе описывается разработка портала знаний, в котором на основе онтологии задач и методов поддержки принятия решений систематизируется имеющаяся в данной области информация, а также предлагаются средства для ее просмотра, анализа и практического использования.

Введение

Разработка систем поддержки принятия решений (СППР) является сложным процессом, требующим усилий специалистов разных типов – экспертов, инженеров знаний, программистов. При этом неизбежно возникает ситуация, когда разработчики вынуждены осваивать смежные специальности. Инженерам знаний приходится стать немного экспертами в той области знаний, для которой создается СППР, и освоить навыки программиста. Эксперты и программисты должны иметь хотя бы общее представление о методах и средствах представления знаний и поддержки принятия решений. При этом в рамках теории принятия решений разработано большое количество методов [Ларичев, 2000], [Петровский, 2009], [Саати, 1993], и в них очень трудно ориентироваться разработчику, не имеющему специальной подготовки. Чтобы успешно решить проблему междисциплинарного взаимодействия, а также повысить скорость и эффективность создания СППР, необходимы общий концептуальный базис и хорошая информационная поддержка, позволяющая разработчику лучше уяснить стоящую перед

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 13-07-00422).

ним задачу и проанализировать доступные методы и средства для ее решения.

В докладе описывается портал знаний по поддержке принятия решений (далее просто портал), который на основе онтологии задач и методов ППР (поддержки принятия решений) систематизирует имеющуюся в данной области информацию, а также предлагает средства для ее просмотра, анализа и практического использования.

1. Содержательные и функциональные компоненты портала

Рассматриваемый портал представляет собой Интернет-ресурс, построенный с использованием технологии разработки порталов научных знаний [Загоруйко и др., 2008] и дополненный новыми возможностями. На рис. 1 показаны его основные содержательные и функциональные компоненты. Знания о предметной области (ПО) портала представлены в виде онтологии, которая помимо понятий и отношений ПО содержит такие концепты, как информационные ресурсы и проблемно-ориентированные сервисы. Информационное наполнение портала, его контент образуют экземпляры онтологии. Таким образом, контент представляет систематизированную в соответствии с онтологией информацию о рассматриваемой ПО.



Рис. 1. Компоненты портала

Вся работа на портале осуществляется с помощью системных сервисов. Сервисы управления онтологией и контентом позволяют создавать новые понятия, отношения и объекты, а также редак-

тировать и удалять ранее созданные. Средства поиска, навигации, фильтрации и визуализации позволяют получать, просматривать и анализировать необходимую пользователю информацию. И, наконец, средства управления проблемно-ориентированными сервисами дают возможность запустить эти сервисы и получить результаты их работы.

Рассмотрим более подробно компоненты портала, обеспечивающие информационную поддержку пользователей.

2. Онтология задач и методов ППР

Качественная информационная поддержка в какой-либо области знаний предполагает, прежде всего, систематизацию понятий данной ПО. Предметной областью рассматриваемого портала является поддержка принятия решений. Концептуальной основой систематизации знаний является онтология, которая строится на основе базовых онтологий научного знания и научной деятельности [Загорулько и др., 2008], а также метаонтологии задач и методов ППР [Загорулько и др., 2012]. Данная метаонтология содержит описание таких базовых понятий поддержки принятия решений, как *Задача*, *Метод*, *Модуль*, *Решатель*, *Входные данные*, *Результат*, *Ситуация*, *Проблемная ситуация*, *Альтернатива*, *Этап принятия решений*, а также отношения между ними.

Онтология задач и методов ППР тесно связана с понятиями предметных областей, для которых создаются СППР. Для того чтобы дать представление о задачах и методах, не вдаваясь в конкретику ПО, описания этих областей должны быть одинаково устроены. Это задается путем использования в качестве основы для построения онтологии конкретной предметной области метаонтологии ПО, включающей базовые понятия, являющиеся общими для всех ПО. Метаонтология ПО включает следующие базовые понятия и отношения: *Объект управления*, *Идеальный объект*, *Состояние*, *Результат*. Объекты могут находиться в том или ином состоянии, для описания которого вводится класс *Состояние*. Результатом работы СППР может быть *Диагноз* текущего состояния объекта (объектов) управления, *Рекомендация* для ЛПР, *Прогноз* изменения состояния объекта и т.п. Отдельным результатом может быть *Задача*, которую необходимо решить, чтобы детально проанализировать объект.

Так как метаонтология ПО определяет только понятия самого верхнего уровня, характерные для всех СППР, независимо от решаемого класса задач, то при специализации СППР на определенную предметную область и для облегчения работы экспертов и инженеров знаний могут разрабатываться онтологии базового уровня, которые затем используются для построения онтологии конкретной ПО.

Важными классами онтологии портала являются проблемно-ориентированные сервисы. Это программные модули, которые реализуют описанные в онтологии методы ППР, имеют стандартизированный интерфейс и могут исполняться по сетевым протоколам.

Помимо систематизированных в онтологии знаний разработчику необходимо структурированное описание конкретных методов. Для этого описания методов даются в соответствии с понятиями онтологии. Их совокупность вместе с существующими между ними отношениями образует семантическую сеть, которая дает представление обо всех аспектах использования конкретных методов в разрабатываемой СППР. Пользователь получит ответы на такие вопросы: «какие задачи решает данный метод», «какие инструментальные методы созданы для его реализации», «какие входные данные необходимы для данного метода», «какие решатели или фреймворки имеются для реализации метода». Помимо этих знаний можно получить как неформальное, так и формализованное описание метода, информацию о группах, коллективах, развивающих данный метод, а также ссылки на Интернет-ресурсы, имеющие отношение к рассматриваемому методу. Для проблемно-ориентированных сервисов дополнительно можно получить ссылку на их реализации.

3. Средства поиска, навигации, фильтрации и визуализации

Качественная информационная поддержка предполагает наличие развитого пользовательского интерфейса. Он должен включать средства поиска требующейся информации, просмотра ее описания. Необходимо также наличие различных аналитических инструментов, подразумевающих такие формы «подачи» и визуализации

просматриваемой информации, которые бы способствовали улучшению ее восприятия.

Имеющиеся на портале сервисы поиска позволяют осуществлять как поиск информации по ключевым словам, так и расширенный семантический поиск с использованием ограничений, задаваемых в терминах онтологии.

Сервис навигации позволяет, просматривая описание понятия или объекта, переходить к описанию связанных с ним понятий и объектов, представляя таким образом его полный контекст.

В качестве аналитических инструментов на портале используются средства фильтрации и визуализации объектов и понятий. Фильтрация позволяет из большого списка выбрать объекты, значения атрибутов которых удовлетворяют указанным ограничениям. Сервис визуализации [Апанович и др., 2011] предоставляет следующие возможности:

1. Отображение полной системы понятий в виде графа (рис. 2, справа).
2. Оценка «мощности» понятий с помощью круговой диаграммы. Под мощностью подразумевается число объектов, относящихся к данному понятию (рис. 2, слева).
3. Отображение полной семантической сети объектов.
4. Отображение сети объектов, связанных выбранным отношением.

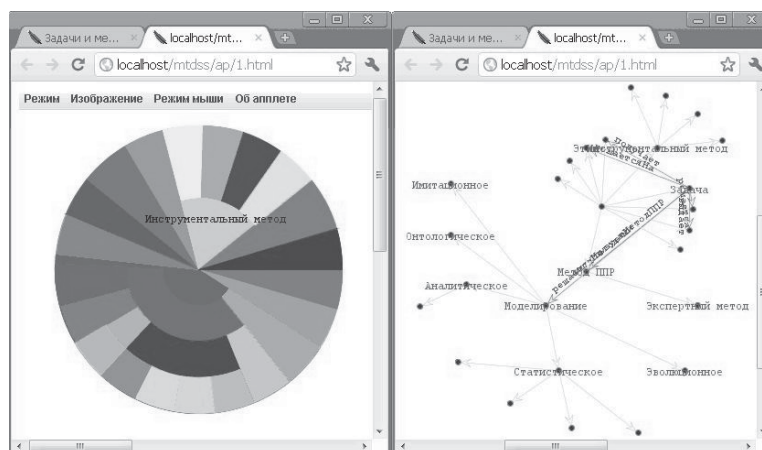


Рис. 2. Средства визуализации

4. Средства управления проблемно-ориентированными сервисами

Средства управления проблемно-ориентированными сервисами позволяют создавать и исполнять эти сервисы непосредственно из портала. Добавление таких средств на портал поднимает возможности оказания информационной поддержки на качественно новый уровень – пользователь не просто получает информацию об интересующем его методе или ссылку на реализацию метода. Он может тут же, на портале, посмотреть примеры использования метода, сам запустить его, проанализировать его работу с разными входными данными.

Для подробного описания методов и реализующих их алгоритмов в интерфейс портала был встроен текстовый редактор. Описания методов, задаваемые или просматриваемые с помощью данного редактора, являются значением атрибута «Подробное описание» соответствующего метода. На рис. 3 представлено описание симплекс-метода для незарегистрированного пользователя. Для зарегистрированных пользователей подробное описание отображается в режиме редактирования с широким набором инструментальных средств, включая средства задания математических формул.

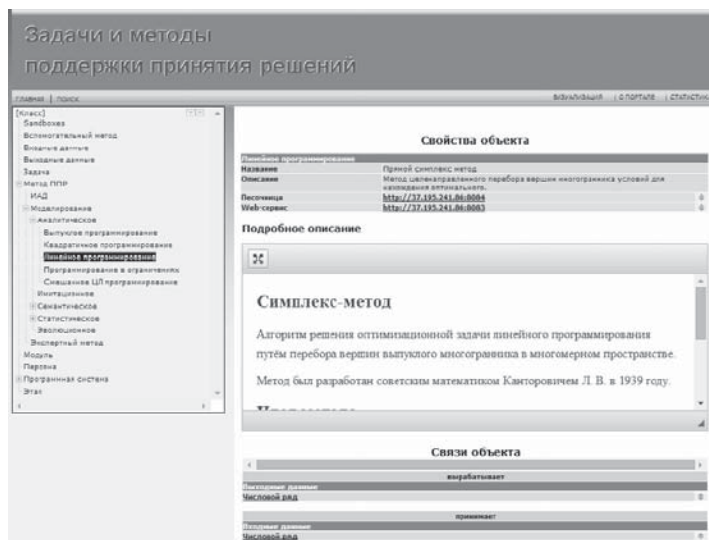


Рис. 3. Пользовательский интерфейс портала знаний по задачам и методам ППР

Для создания проблемно-ориентированных сервисов был разработан Веб-сервис, который на данном этапе реализует методы математического программирования. Веб-сервис снабжен пользовательским интерфейсом, позволяющим просматривать спецификации реализуемых им методов.

Возможность практического ознакомления с работой методов поддержки принятия решений реализована в виде веб-ресурса, являющегося своего рода «песочницей», позволяющей задать входные данные, запустить исполнение метода и посмотреть результаты его работы.

5. Реализация средств управления сервисами

При разработке данного портала возможности, определяемые используемой технологией (см. п. 1), были расширены описанными выше средствами.

Сторонний текстовый редактор для задания подробного описания SKeditor написан на языке JavaScript и имеет богатый программный интерфейс для работы с ним. К данному редактору был добавлен плагин, позволяющий редактировать математические формулы в формате LaTeX.

Для разработки веб-сервиса, реализующего методы математического программирования, была выбрана платформа ASP.NET [Мак-Дональд и др., 2011] как самая распространенная платформа для создания подобных ресурсов. Языком программирования был выбран C# как основной язык платформы .NET. Данная платформа подразумевает создание веб-сервиса с использованием протокола передачи данных SOAP, работающего поверх протокола HTTP, и языка WSDL.

Для реализации конкретных методов веб-сервиса была выбрана библиотека математического программирования, моделирования и оптимизации Microsoft Solver Foundation. Эта библиотека также основана на платформе .NET и позволяет строить и решать модели задач математического программирования, написанные на специальном языке Optimization Modeling Language (OML).

Веб-ресурс, предназначенный для апробации методов, реализованных в веб-сервисе, разработан на платформе ASP.NET WebForms и языка C#.

Заключение

В докладе описан портал знаний, предназначенный для информационно-аналитической поддержки разработчиков СППР, концептуальной основой которой является онтология, базирующаяся на онтологиях научного знания, научной деятельности и метаонтологии задач и методов ППР.

При построении портала была использована технология разработки порталов научных знаний и добавлены новые функциональные возможности.

Для подробного описания методов был добавлен текстовый редактор с расширением, позволяющим редактировать математические формулы. Был разработан веб-сервис, реализующий методы решения задач линейного, выпуклого и квадратичного программирования. Для апробации методов, реализованных в веб-сервисе, был разработан веб-ресурс, позволяющий задать входные данные и получить результат работы метода.

Расширенный новыми возможностями портал является удобным средством информационной поддержки разработчиков СППР. Он не только предоставляет пользователю систематизированную информацию о разных аспектах построения СППР, но также позволяет опробовать методы ППР, которые далее будут использованы в разрабатываемых системах.

Дальнейшее развитие портала по поддержке принятия решений будет происходить в следующих направлениях:

1. Перенос портала с платформы языка PHP на более современную и легко масштабируемую ASP.NET MVC + C#.
2. Дальнейшее наполнение базы знаний портала и увеличение его функциональности.
3. Наполнение веб-сервиса новыми методами для решения задач ППР.
4. Добавление возможности композиции нескольких сервисов.

СПИСОК ЛИТЕРАТУРЫ

[Апанович и др., 2011] Апанович З.В., Винокуров П.С., Кислицина Т.А. Методы и средства визуализации информационного наполнения больших научных порталов // Вестник НГУ Серия: Информационные технологии. 2011. – Том 9, выпуск 3. Редакционно-издательский центр НГУ. – С. 5–14.

[Глоба и др., 2012] Глоба Л.С., Новогрудская Р.Л. Модели и методы интеграции информационных и вычислительных ресурсов //Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2012): материалы II Междунар. научн.-техн. конф. (Минск, 16–18 февраля 2012 г.) / редкол.: В.В. Голенков (отв. ред.). – Минск: БГУИР, 2012. – С. 447–452.

[Загорулько и др., 2008] Загорулько Ю.А., Боровикова О.И. Подход к построению порталов научных знаний // Автометрия. 2008. Т. 44. № 1. С. 100–110.

[Загорулько и др., 2012] Загорулько Г.Б., Загорулько Ю.А. Подход к разработке онтологии задач и методов поддержки принятия решений // Труды 13-й национальной конференции по искусственному интеллекту с международным участием КИИ-2012. – Белгород: Изд-во БГТУ, 2012. – Т.2. – С. 185–192.

[Ларичев, 2000] Ларичев О.И. Теория и методы принятия решений, а также хроника событий в Волшебных странах. – М.: Логос, 2000.

[Мак-Дональд и др., 2011] Мак-Дональд М., Фримен А., Шпушта М. Microsoft ASP.NET 4 с примерами на C# 2010 для профессионалов, 4-е изд.: Пер. с англ. – М.: ООО «И. Д. Вильямс», 2011. – 1424 с.

[Петровский, 2009] Петровский А.Б. Теория принятия решений. – М.: Издательский центр «Академия», 2009.

[Саати, 1993] Саати Т. Л. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993.

[Феррара и др., 2003] Феррара А., Мак-Дональд М. Программирование web-сервисов для .NET. Библиотека программиста СПб.: Питер, 2003. – 430 с.

УДК 002.53:004.89

ПОДХОД К АВТОМАТИЗАЦИИ СБОРА ИНФОРМАЦИИ ДЛЯ ТЕМАТИЧЕСКИХ ИНТЕЛЛЕКТУАЛЬНЫХ НАУЧНЫХ ИНТЕРНЕТ-РЕСУРСОВ¹

Ю.А. Загорулько (*zagor@iis.nsk.su*)

О.И. Боровикова (*olesya@iis.nsk.su*)

Е.А. Сидорова (*lena@iis.nsk.su*)

Институт систем информатики имени А.П. Ершова СО РАН,
Новосибирск

И.Р. Ахмадеева (*ah.irishka@gmail.com*)

Новосибирский государственный университет, Новосибирск

В работе представлен подход к автоматизации сбора информации для тематических интеллектуальных научных интернет-ресурсов, обеспечивающих систематизацию и интеграцию научных знаний, информационных ресурсов, относящихся к определенной области знаний, и средств их интеллектуальной обработки, а также содержательный доступ к ним. Подход основан на использовании онтологии и тезауруса моделируемой области знаний.

Введение

В мире накоплено огромное количество информации по различным областям знаний, причем значительная ее часть представлена непосредственно в сети Интернет, но, несмотря на это, проблема эффективного обеспечения научного сообщества информацией по интересующим его тематикам остается на повестке дня. Нерешенной остается и проблема удобного доступа к средствам обработки данных, собранных по этим тематикам. Даже уже представленные в Интернет в виде веб-сервисов реализации методов их обработки остаются недоступными широкому кругу пользователей из-за отсутствия содержательной информации о них.

Для решения этих проблем, в частности для информационной и аналитической поддержки научной и производственной деятельности в определенных областях знаний, создаются тематические

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 13-07-00422).

интеллектуальные научные интернет-ресурсы (ИНИР) [Загорюлько и др., 2013].

Сбор и накопление такой информации в ИНИР – довольно трудоемкая задача, требующая больших усилий многих специалистов. Ввиду высокой потребности в системах такого класса с целью ускорения их создания и облегчения сопровождения необходимо автоматизировать сбор релевантной информации по тематике ИНИР из сети Интернет. Описанию подхода, поддерживающего такую автоматизацию, и посвящена данная работа.

Следует заметить, что проблемой сбора информации из Интернет занимаются многие разработчики и исследователи. Однако, как показывает обзор [Ferrara al., 2014], большая часть исследований направлена на извлечение информации, необходимой для решения задач электронной коммерции или анализа новостного потока и социальных сетей, и только их незначительная часть – на извлечение информации для нужд научной деятельности [DeRose et al., 2007; Ланин и др., 2010].

1. Концепция и система знаний ИНИР

Ядро информационной модели ИНИР составляет онтология, которая, вводя формальные описания понятий некоторой области знаний, типов информационных ресурсов и методов их интеллектуальной обработки в виде классов объектов и отношений между ними, одновременно задает структуры для представления информации о реальных объектах моделируемой области знаний, интегрируемых информационных ресурсах и методах и средствах их обработки. Данная информация хранится в контенте ИНИР в виде семантической сети, типы информационных объектов и отношений которой определяются классами объектов и отношений онтологии ИНИР.

На основе онтологии организуется удобная навигация по научным знаниям и информационным ресурсам, интегрированным в ИНИР, а также содержательный поиск данных и средств их интеллектуальной обработки.

В систему знаний ИНИР включен также тезаурус, содержащий термины моделируемой области знаний, т.е. слова и словосочетания, с помощью которых понятия онтологии представляются в текстах и пользовательских запросах. Тезаурус задает смысл понятий

посредством соотнесения одних понятий с другими с помощью семантических отношений. Благодаря этому он может применяться при поиске и аннотировании информационных ресурсов, интегрируемых в ИНИР.

Поскольку предлагаемый подход к сбору информации из сети Интернет существенно базируется на онтологии ИНИР, рассмотрим ее подробнее.

Онтология ИНИР состоит из трех взаимосвязанных онтологий: онтологии области знаний ИНИР, онтологии научных информационных ресурсов и онтологии задач и методов.

Онтология области знаний ИНИР строится на основе двух базовых онтологий – онтологии научной деятельности и онтологии научного знания. Первая из этих онтологий включает классы понятий, относящиеся к организации научной и исследовательской деятельности, такие как *Персона*, *Организация*, *Событие*, *Конференция*, *Проект*, *Публикация* и др. Вторая базовая онтология содержит понятия, необходимые для представления научных дисциплин. В частности, она содержит такие классы, как *Раздел науки*, *Метод исследования*, *Объект исследования*, *Научный результат* и др.

Онтология задач и методов предназначена для описания задач, на решение которых нацелен ИНИР, и методов их решения, включая методы интеллектуальной обработки данных, представленных в информационных ресурсах.

Основным классом онтологии научных информационных ресурсов является класс *Информационный ресурс*, служащий для описания информационных ресурсов. Набор атрибутов и связей этого класса основан на стандарте Dublin core [Hillmann, 2005]. Он имеет следующие атрибуты: *название ресурса*, *язык ресурса*, *тематика ресурса*, *тип доступа к ресурсу* и т.п. Объекты этого класса связываются семантическими отношениями с другими информационными объектами, представляющими в контенте ИНИР организации, персоны, публикации, проекты, разделы науки и т.п.

В этом классе выделяются два крупных подкласса – *Информационный интернет-ресурс* и *Web-сервис*. Первый из них служит для представления таких интернет-ресурсов, как сайты организаций, конференций, научно-исследовательских проектов и т.п. Второй – для представления web-сервисов, реализующих различные методы обработки информации.

2. Модель сбора информации для ИНИР

Сложность задачи сбора информации для ИНИР определяется большим разнообразием видов извлекаемой информации и способов ее представления в Интернет. В частности, необходимо собирать информацию об организациях, проектах, публикациях, интернет-ресурсах, веб-сервисах и других сущностях, описываемых онтологией научной деятельности. Эта информация может быть представлена как в виде интернет-страниц, имеющих различную структуру, так и в виде текстовых документов в различных форматах. В связи с этим мы посчитали нецелесообразным использовать популярные в настоящее время методы извлечения информации, основанные на обучении на примерах (см. например, [Zhai et al., 2005]), а применили подход, базирующийся на онтологии. В соответствии с этим было решено для каждого типа сущностей (класса онтологии научной деятельности) разработать свой метод сбора и обработки информации, настраиваемый на предметную область и типы интернет-ресурсов и документов.

Заметим, что предлагаемый подход развивает методы сбора онтологической информации об интернет-ресурсах [Загоруйко, 2008], разработанные в рамках технологии построения порталов научных знаний. В то же время он близок к подходу, представленному в [DeRose et al., 2007], который базируется на концептуальной модели предметной области и предлагает использовать для каждого вида сущности свои шаблоны и обработчики.

Сбор информации для ИНИР включает следующие этапы:

1. Поиск релевантных областей знаний ИНИР интернет-ресурсов и документов.
2. Извлечение информации из найденных интернет-ресурсов и документов.
3. Занесение полученной информации в контент ИНИР.

В соответствии с этим подсистема сбора информации из сети Интернет включает модуль поиска релевантных интернет-ресурсов, модуль извлечения информации из интернет-ресурсов, модуль занесения найденной информации в контент ИНИР, а также базу данных ссылок на интернет-ресурсы (БД СИР).

2.1. Поиск релевантных интернет-ресурсов и документов

При настройке ИНИР на область знаний выполняется заполнение БД СИР ссылками на релевантные, по мнению экспертов,

интернет-ресурсы. При этом для каждой ссылки указывается класс онтологии, объекты которого описывает соответствующий ей ресурс. С каждой ссылкой также связывается следующая метаинформация: дата загрузки, частота обновления, периодичность повторной закладки, дата последней проверки, статус обработки. Первые четыре параметра вводятся для отслеживания актуальности ресурса, последний – для указания статуса ссылки (релевантная, нерелевантная, необработанная).

Список ссылок пополняется не только вручную, но и автоматически – модулем поиска интернет-ресурсов, который выполняет сбор ссылок на релевантные интернет-ресурсы по поисковым запросам, сформированным на основе названий классов онтологии и терминов тезауруса, представляющих понятия моделируемой области знаний. Он запускается с заданной при настройке ИНИР периодичностью. При этом модуль поиска обращается к поисковым системам Google, Яндекс и Bing через их программные интерфейсы, т.е. использует механизм метапоиска с последующей фильтрацией дубликатов и нерелевантных ссылок [Ахмадеева и др., 2013].

БД СИР может также пополняться ссылками, обнаруженными при извлечении информации из обрабатываемых интернет-ресурсов. Эти ссылки в дальнейшем анализируются экспертами, которые принимают решение об их релевантности.

2.2. Извлечение информации

Для заполнения контента ИНИР собирается информация из таких источников, как порталы знаний, электронные библиотеки и журналы, сайты организаций, ассоциаций, проектов и конференций, новостные ленты, социальные научные сети, вики-ресурсы, реестры (каталоги) веб-сервисов и др. Как было сказано выше, из этих источников извлекается информация о проектах, организациях, персонах, конференциях и публикациях, т.е. обо всех объектах базовых классов онтологии научной деятельности, а также об объектах классов *Информационный интернет-ресурс* и *Web-сервис* онтологии научных информационных ресурсов.

Для каждого из этих классов создается свой метод извлечения информации, включающий набор шаблонов. В шаблонах для каждого типа извлекаемой информации указываются обработчики, реализующие алгоритмы обхода и анализа соответствующих фраг-

ментов интернет-страниц или документов. Указанные шаблоны генерируются на основе онтологии. Для повышения полноты извлечения информации увеличивается вариативность этих шаблонов за счет использования в них альтернативных терминов из тезауруса (синонимов и гипонимов).

Модуль извлечения информации осуществляет анализ интернет-ресурсов, которые он скачивает по ссылкам, заданным в БД СИР.

Для облегчения анализа HTML-страниц ресурса представляется в виде DOM-дерева в соответствии со стандартом DOM (Document Object Model), регламентирующим способ представления содержимого документа (в частности, HTML-страницы) в виде набора объектов [Stenback et al., 2003].

Анализ DOM-дерева каждой страницы выполняется на основе соответствующего шаблона, при этом определяется релевантность загруженной страницы тематике ИНИР и извлечение описанной этим шаблоном информации.

Например, интернет-ресурс, на котором размещена информация о проекте, может быть представлен сайтом проекта, разделом сайта организации или персоны или публикацией, описывающей проект. Для каждого из этих способов представления на основе класса онтологии *Проект* строится свой шаблон.

Рассмотрим один из вариантов организации такого шаблона, но сначала дадим описание свойств класса *Проект*.

```
class Проект (Название: string; Аббревиатура: string; Описание:
string;
    Дата начала: date; Дата окончания: date; Номер: string;
URL: string;
    Стадия: Этап_проекта; Ключевые слова: set_of_string)
relation Проект_Включает < Проект, Проект >
relation Проект_Поддерживается < Проект, Организация >
relation Задача_Проекта < Проект, Задача >
relation Участник_Проекта < Проект, Персона > (Роль: set_of_
Роль)
relation Участник_Проекта_Орг < Проект, Организация >
relation Научное_направление < Проект, Раздел_науки >
relation Результат_Деятельности < Проект, Результат >
relation Исследует_Объект < Проект, Объект_исследования >
relation Использует_Метод < Проект, Метод_исследования >
```

relation Публикация_о_Проекте < Проект, Публикация >
relation Интернет_Ресурс_Проекта < Проект, Интернет_Ресурс >

Шаблон для извлечения информации будет иметь следующий вид:

```
<Class Name= «Проект» engine = FragmentSearch >  
  < Marker Term = «О проекте» PType=Menu FragType=Page />  
  < Marker Term = «Проект» PType=Head FragType=Block />  
  <Attr Name= «Название» type= string engine = NameEntity >  
    <Marker Term=«Проект» PType=link FragType=LinkText/>  
    <Marker Term=«Проект» PType=sentence  
    FragType=QuoteText/>  
    <Marker Term= «Проект» PType=Head FragType = Head />  
  </Attr>  
  < Attr Name= «Аннотация» type=text >  
    <Marker Term= «Аннотация» PType= Head FragType=Block />  
    <Marker Term = «О проекте» PType=Menu FragType=Page />  
  </Attr>  
<Relation Name = «Публикация_о_Проекте» >  
  <Marker Term = «Публикации» PType= Menu FragType=Page />  
  <Marker Term = «Литература» PType= Head FragType=Block />  
  <Object Name = «Публикация» engine = PublicationList />  
</Relation>  
<Relation Name= «Участник проекта» >  
  <Marker Term= «Об участниках» PType=Menu FragType=Page />  
  <Marker Term= «Список участников» PType= Head  
  FragType=Block/>  
  <Object Name= «Персона» engine = PersonList />  
</Relation> </Class>
```

Шаблон, предназначенный для извлечения объектов заданного класса, описывается блоком *Class* и содержит блоки атрибутов (*Attr*), отношений (*Relation*) и аргументов отношений (*Object*). Каждый из этих блоков может описываться одним или группой альтернативных маркеров (*Marker*), задающих свойства фрагмента текста, содержащего извлекаемую информацию. Маркер, приписанный непосредственно блоку *Class*, выделяет текстовый фрагмент, описывающий объект и определяющий область дальнейшего поиска маркеров.

К параметрам маркера относятся: (1) Term – термин тезауруса, представленный множеством альтернативных написаний термина, (2) RType – тип фрагмента, в тексте которого должен располагаться термин маркера, (3) FragType – тип фрагмента, который должен извлекаться, (4) engine – имя обработчика, который будет извлекать необходимую информацию в найденном по маркеру фрагменте.

Анализ входной страницы, представленной после предварительной обработки в структурированном виде (DOM-структура), осуществляется обработчиком верхнего уровня, который решает следующие задачи:

- поиск шаблона, подходящего для данной страницы или ее фрагмента, на основе маркеров блока Class;
- поиск маркерных терминов и извлечение текстовых фрагментов в соответствии с параметрами маркера;
- вызов специализированных обработчиков, формирование входных данных и обработка результата их работы;
- формирование объекта заданного онтологического класса и его связей.

Например, на сайте проекта «Национальный корпус русского языка» (<http://www.ruscorgo.ru>) в разделе меню «О проекте» можно найти краткое описание проекта, в разделе «участники проекта» – информацию о персонах и организациях, участвующих в проекте, в разделе «публикации» – информацию о публикациях по теме проекта и т.д.

Шаблон, построенный на основе класса *Проект*, позволит извлечь эту информацию со страниц данного сайта. При этом для извлечения информации, составляющей контекст проекта и, как правило, определяемой отношениями класса *Проект*, например, данных о публикациях по теме проекта, персонах и организациях, участвующих в проекте, используются обработчики и шаблоны, специально построенные для извлечения информации такого типа и многократно используемые в других шаблонах, соответствующих базовым понятиям онтологии.

Заключение

ИНИР позволяет исследователям значительно сократить время, требуемое для обеспечения доступа к необходимой информации и ее анализа, за счет аккумуляции описаний релевантных интернет-

ресурсов и методов их обработки. При этом использование ИНИР будет тем эффективнее, чем более полно в нем будет представлена информация по его тематике. Добиться такой полноты можно только за счет автоматизации сбора информации из сети Интернет.

В настоящее время реализован ряд компонентов подсистемы сбора информации из сети Интернет, а именно: модуль поиска релевантных интернет-ресурсов, модуль извлечения информации, база данных ссылок на интернет-ресурсы. На данный момент разработан метод извлечения данных о проекте, включая сопутствующие ему шаблоны и обработчики, реализующие извлечение информации о персонах и публикациях. Заметим, что для анализа списков публикаций и персон в описаниях проектов используются ранее разработанные нами средства генерации формальных описаний научных статей [Дяченко, 2010].

СПИСОК ЛИТЕРАТУРЫ

[Ахмадеева и др., 2013] Ахмадеева И.Р., Загорулько Ю.А., Саломатина Н.В., Серый А.С., Сидорова Е.А., Шестаков В.К. Подход к формированию тематических коллекций текстов на основе интернет-ресурсов // Вестник НГУ. Серия: Информационные технологии. 2013. Том.11, выпуск 4.

[Дяченко, 2010] Дяченко О.О., Загорулько Ю.А. Генерация формальных описаний научных статей для информационных систем // Труды 12-й национальной конференции по искусственному интеллекту с международным участием – КИИ-2010. – М.: Физматлит, 2010.

[Загорулько, 2008] Загорулько Ю.А. Автоматизация сбора онтологической информации об интернет-ресурсах для портала научных знаний // Известия Томского политехнического университета. – Т. 312. – № 5. Управление, вычислительная техника и информатика. – 2008.

[Загорулько и др., 2013] Загорулько Ю.А., Загорулько Г.Б., Шестаков В.К., Кононенко И.С. Концепция и архитектура тематического интеллектуального научного интернет-ресурса // Труды XV Всероссийской научной конференции RCDL'2013, Ярославль, Россия, 2013.

[Ланин и др., 2010] Ланин В.В., Мальцев П.А., Лядова Л.Н. Технологии сбора и анализа информации для исследовательского портала // Материалы четвертой Международной научно-технической конференции «Инфокоммуникационные технологии в науке, производстве и образовании» (Инфоком 4): Часть I, 2010.

[DeRose et al., 2007] DeRose P., Shen W., Chen F., Doan AH, Ramakrishnan R. Building Structured Web Community Portals: A Top-Down, Compositional, and Incremental Approach // VLDB '07, September 23–28, 2007, Vienna, Austria.

[Ferrara et al., 2014] Ferrara E., De Meo P., Fiumara G., Baumgartner R.. Web Data Extraction, Applications and Techniques: A Survey // Preprint submitted to Knowledge-based systems. June 5, 2014.

[Hillmann, 2005] Hillmann D. Using Dublin Core, 2005. – <http://dublin-core.org/documents/usageguide/>

[Stenback et al., 2003] Stenback J., Le Hégaré P., Le Hors A. Document Object Model (DOM) Level 2 HTML Specification // W3C Recommendation, 2003. – <http://www.w3.org/TR/2003/REC-DOM-Level-2-HTML-20030109/>

[Zhai et al., 2005] Zhai Y., Liu B. Extracting Web Data Using Instance-Based Learning. // Proceedings of 6th Inter-national Conference on Web Information Systems Engineering (WISE-05), 2005.

УДК 004.056

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА МОНИТОРИНГА И УПРАВЛЕНИЯ ИНЦИДЕНТАМИ КИБЕРБЕЗОПАСНОСТИ¹

И.В. Котенко (*ivkote@comsec.spb.ru*)

И.Б. Саенко (*ibsaen@comsec.spb.ru*)

Санкт-Петербургский институт информатики и автоматизации РАН,
Санкт-Петербург

В работе рассматриваются методы и архитектура интеллектуальной системы мониторинга и управления инцидентами кибербезопасности, расширяющей возможности традиционных систем такого типа. Приведено формальное описание методов, используемых на начальных этапах мониторинга и для принятия решений, дан пример их реализации, рассмотрены реализованная архитектура системы и области ее тестирования.

Введение

Мониторинг и управление инцидентами кибербезопасности являются важнейшей составляющей защиты информации в киберпространстве. Системы, реализующие данную функцию, известны как SIEM-системы (Security Information and Event Management) [Miller et al., 2011]. Их работа основана на сборе, хранении и последующем использовании для аналитической обработки сведений о событиях безопасности, которые фиксируются в системных журналах различных источников: операционных систем, систем управления базами данных, антивирусов, систем обнаружения кибератак, сетевых элементов [Котенко и др., 2012].

В настоящее время проводятся многочисленные исследования по расширению функциональных возможностей систем такого класса. Интеллектуализация является основным направлением такого развития, обеспечивающим возможность эффективно использовать SIEM-системы в компьютерных сетях большой размерности, объединять обработку событий сетевого, физического и бизнес

¹ Работа выполнена при финансовой поддержке РФФИ (13-01-00843, 13-07-13159, 14-07-00697, 14-07-00417), программы фундаментальных исследований ОНИТ РАН (контракт №2.2) и проекта ENGENSEC программы Европейского Сообщества TEMPUS.

уровней и иметь высокую достоверность в условиях неполноты и противоречивости исходных данных [Котенко и др., 2012].

Несмотря на то, что известны многие работы, посвященные интеллектуальным методам защиты информации, например, [Варламов, 2006], [Галатенко, 2007], работы, в которых дается единое описание этих методов с формальных позиций, практически отсутствуют, а такое описание, на наш взгляд, является неотъемлемым условием дальнейшего успешного развития систем данного класса.

1. Формальное описание интеллектуальной системы мониторинга и управления инцидентами безопасности

Интеллектуальная система мониторинга и управления инцидентами кибербезопасности может быть представлена в следующем виде: $IM_SMM = \langle Repository, Met_Search, Met_PreProc, Met_Analys \rangle$, где *Repository* – репозиторий данных и знаний, являющийся системообразующим компонентом; *Met_Search* – методы сбора данных; *Met_PreProc* – методы предварительной обработки; *Met_Analys* – методы анализа.

Репозиторий представлен в виде: $Repository = \langle MetaDat, Dat \rangle$, где *MetaDat* – метаданные (схема базы знаний); *Dat* – экстенциональные данные (содержимое базы знаний). Метаданные: $MetaDat = \langle Concepts, Relations, Actions, Rules \rangle$, где *Concepts* – множество концептов; $Relations = \langle C1, C2, Type \rangle$ – множество связей между концептами (*C1* и *C2*), относящихся к некоторому множеству *Type* типов связей; *Actions* – множество действий, которые следует предпринимать, если выполняются правила, составляющие множество *Rules*. Экстенциональные данные репозитория: $Dat = \langle Attr_Concept, Attr_Relation \rangle$, где *Attr_Concept* – множество атрибутов, характеризующих экземпляры концептов, *Attr_Relation* – множество атрибутов, характеризующих экземпляры связей.

Методы сбора данных состоят из двух типов методов: $Met_Search = \langle Met_Search_Out, Met_Search_In \rangle$, где *Met_Search_Out* – метод сбора данных, инициируемый внешним источником данных; *Met_Search_In* – метод сбора данных, инициированный самой системой. Методы предварительной обработки данных: $Met_PreProc = \langle Met_Norm, Met_Filtr, Met_Aggr, Met_Corr, Met_Prior \rangle$, где *Met_Norm* – метод нормализации; *Met_Filtr* – метод фильтрации; *Met_Aggr* – метод агрегации; *Met_Corr* – метод корреляции; *Met_Prior* – метод приоритизации. Методы анализа: $Met_Analys =$

$\langle Met_Eval, Met_Predict, Met_Decision \rangle$, где Met_Eval – метод оценки уровня кибербезопасности; $Met_Predict$ – метод прогнозирования кибербезопасности; $Met_Decision$ – метод принятия решения по обеспечению кибербезопасности. Ниже подробнее рассмотрим методы первых двух групп и метод принятия решения на основе онтологического подхода. Остальные методы анализа требуют отдельного рассмотрения.

2. Методы сбора данных

2.1. Метод сбора, инициируемый внешним источником

Сущность метода заключается в определении момента времени, в который необходимо направить данные о событии безопасности в центральную часть системы. Естественно полагать, что таким моментом является момент возникновения события безопасности. Целесообразно передавать сгруппированные события в соответствии с установленным расписанием. Расписание представим в виде следующего временного ряда: $Schedule = \{(Time_i, Interval_i)\}$, где $Time_i$ – i -ый момент времени; $Interval_i$ – интервал, через который следует передавать пакеты связи для i -ого момента времени. Условие включения данных о событии безопасности, произошедшего в момент времени $Time^*$, в пакет связи, подлежащий отправке в момент времени $Time_0$, имеет следующий вид: $((Time_0 - Time^*) < Interval_i) \text{ AND } (Time_0 \in [Time_i, Time_{i+1}]) \text{ AND } (Severity^* < Severity_0)$, где $[Time_i, Time_{i+1}]$ – интервал времени в расписании $Schedule$, ограниченный моментами времени $Time_i$ и $Time_{i+1}$; $Severity^*$ – степени важности события; $Severity_0$ – степень важности события пороговая.

Если справедливо $Severity^* > Severity_0$, то осуществляется незамедлительная передача пакета связи. Поэтому метод включает следующие шаги: (1) фиксация времени возникновения события $Time^*$ и его важности $Severity^*$; (2) сравнение важности $Severity^*$ с пороговым значением $Severity_0$. Если оно его превышает, то формирование пакета связи, включающего только одно возникшее событие и переход на шаг 6; (3) определение интервала расписания $[Time_i, Time_{i+1}]$ по условию $Time_i < Time^* < Time_{i+1}$; (4) определение интервала между пакетами связи $Interval_i$, соответствующего моменту $Time_i$; (5) включение события в пакет связи, для которого выполняется условие $(Time_0 - Time^*) < Interval_i$; (6) передача пакета связи.

2.2. Метод сбора, инициируемый центральной частью системы

Данный метод отличается от предыдущего тем, что необходимость формирования и передачи пакета связи инициируется системой на основании анализа данных, содержащихся в репозитории.

Условие, согласно которому формируется команда источнику данных на передачу данных о событиях безопасности, имеет вид: $\exists Event_j, (Severity (Event_j | Met_Eval) > Severity_0) \text{ AND } (Location (Event_j) = Location_0) \text{ AND } (Time (Event_j | Met_Predict) < Time_{i+1})$, где $Event_j$ – событие, хранящееся в репозитории; $Severity (Event_j | Met_Eval)$ – критичность события $Event_j$, рассчитанное с помощью метода Met_Eval ; $Location (Event_j)$ – местоположение события $Event_j$; $Location_0$ – местоположение источника данных; $Time (Event_j | Met_Predict)$ – время наступления события $Event_j$, рассчитанное с помощью метода $Met_Predict$.

Метод включает следующие шаги: (1) определение критичности события $Severity (Event_j | Met_Eval)$; (2) если критичность ниже, чем $Severity_0$, то – останов; (3) определение местоположения $Location (Event_j)$; (4) определение времени $Time (Event_j | Met_Predict)$; (5) передача запроса источнику с местоположением $Location_0$ на пересылку данных о событиях безопасности в интервале времени $Interval_i$.

3. Методы предварительной обработки данных

3.1. Метод нормализации событий безопасности

Сущность данного метода заключается в приведении данных к базовому системному формату. В качестве такового предлагается выбрать один из стандартов представления информации и событий безопасности – СВЕ [Bridgewater, 2004], СЕЕ [Common, 2008] или XDAS [Open, 2011]. Положим, $Format_0$ – базовый формат, $Format_j$ – формат j -го источника, $Convert (E, F)$ – операция преобразования данных о событии безопасности E в формат F . Тогда метод записывается в виде:

$$\exists Event_j, Format_j \neq Format_0 \rightarrow Convert (Event_j, Format_0).$$

3.2. Метод фильтрации событий безопасности

Сущность данного метода заключается в исключении из входного потока данных повторных или дублируемых событий. Для

его реализации необходима вспомогательная буферная память, в которой хранятся инциденты безопасности: $Store_Ins = \{inc_filtr_k\}$, где $inc_filtr_k = \langle attr_k, val_attr_k \rangle$, $attr_k \subset MetaDat$, $val_attr_k \subset Dat$ и val_attr_k – значения, которые принимают атрибуты $attr_k$ инцидента inc_filtr_k . Под инцидентом безопасности понимается «любое событие, не являющееся частью стандартных операций по предоставлению услуги, которое привело или может привести к нарушению или снижению качества этой услуги» [ССТА, 2000]. Условием исключения события из входного потока является тот факт, что в буферной памяти уже имеется инцидент, описывающий данное событие.

В методе фильтрации используются две операции обработки инцидентов: $Create(inc_filtr)$ – создание нового инцидента; $Compare(E, inc_filtr)$ – сравнение события E с хранимым инцидентом inc_filtr . Последняя операция принимает значение $True$ или $False$ и используется для непосредственной фильтрации событий.

Операция создания нового инцидента записывается в следующем виде:

$$\exists Event_j, \forall inc_filtr_k \in Store_Ins, (attr(Event_j) \neq attr_k) \text{ OR } (val_attr(Event_j) \neq val_attr_k) \rightarrow \text{создать } inc_filtr_{k0} \in Store_Ins, (attr_{k0} = attr(Event_j)) \text{ AND } (val_attr_{k0} = val_attr(Event_j)).$$

Операция сравнения записывается следующим образом:

$$\exists Event_j, \forall inc_filtr_k \in Store_Ins, (attr(Event_j) \neq attr_k) \text{ OR } (val_attr(Event_j) \neq val_attr_k), Compare = True \rightarrow Compare = False.$$

Тогда метод фильтрации имеет следующее описание:

$$\exists Event_j, \forall inc_filtr_k \in Store_Ins, Compare(Event_j, inc_filtr_k) = False \rightarrow Create(inc_filtr_{k0}),$$

$$\exists Event_j, \forall inc_filtr_k \in Store_Ins, Compare(Event_j, inc_filtr_k) = True \rightarrow \text{удалить } Event_j \text{ из входящего потока.}$$

3.3. Метод агрегации событий безопасности

Сущность данного метода заключается в объединении нескольких событий входного потока данных, имеющих схожие атрибуты. Чтобы реализовать этот метод, вводится операция $Similar(E, inc)$ определения подобности события E хранимому инциденту inc и операция $Aggr(E_1, E_2)$ агрегирования событий E_1 и E_2 . Операция $Similar$ имеет следующий вид: $\exists Event_j, \forall inc_k \in Store_Ins, (attr(Event_j) \subseteq attr_k) \text{ AND } (val_attr(Event_j) \subseteq val_attr_k), Similar = False \rightarrow Similar = True$. Операция агрегирования: $\exists E_1, E_2, Similar(E_1, E_2)$

$= True, Similar (E_2, E_1) = True \rightarrow Aggr = E_3, attr (E_3) = attr (E_1) \cup attr (E_2), val_attr (E_3) = val_attr (E_1) \cup val_attr (E_2)$. С помощью введенных операций метод агрегации событий безопасности записывается в следующем виде: $\exists E_1, E_2, \exists inc_k \in Store_Ins, Similar (E_1, inc_k) = True, Similar (E_2, inc_k) = True \rightarrow Aggr (E_1, E_2)$.

3.4. Методы корреляции событий безопасности

Сущность данных методов заключается в выявлении во входном потоке данных инцидентов на основании сигнатур либо правил, хранящихся в памяти. Условие, по которому принимается решение о выявлении инцидента сигнатурным методом, имеет вид: $\exists Event_j, \exists inc_k \in Store_Ins, sign_k = Signature (inc_k), (attr (sign_k) \subseteq attr (Event_j)) \text{ AND } (val (sign_k) \subseteq val_attr (Event_j)) \rightarrow Event_j = inc_k$, где $attr (sign_k)$ – множество атрибутов сигнатуры; $val (sign_k)$ – множество значений, которые принимают атрибуты этой сигнатуры.

Правило, используемое для корреляции, имеет вид $Rule_m: E_1, E_2, \dots, E_n \rightarrow E_{n+1}$. Выявление инцидента inc_k согласно методу, основанному на правилах, происходит на основании выполнения множества приписанных данному инциденту правил: $inc_k \leftarrow Rule_{1k} (Rule_{2k} (\dots (Rule_{mk})))$. Поэтому данный метод может быть записан в следующем виде: $\exists Event_j, \exists inc_k \in Store_Ins, Rule (inc_k) = \{Rule_{1k}, Rule_{2k}, \dots, Rule_{mk}\}, Event_j = Rule_{1k} (Rule_{2k} (\dots (Rule_{mk}))) \rightarrow Event_j = inc_k$.

3.5. Метод приоритизации событий безопасности

Сущность данного метода заключается в определении значимости событий во входном потоке данных, которая зависит от степени угрозы. Каждый инцидент inc_k характеризуется множеством признаков $Features (inc_k) = \{feat_{kr}\}$ и множеством угроз $Threat (inc_k) = \{thr_{ks}\}$, каждая из которых оценивается степенью угрозы $Severity (thr_{ks})$. Используя введенные понятия, данный метод может быть записан в виде: $\exists Event_j, \exists feat_i \in Store_Ins, \exists inc_k \in Store_Ins, feat_i \in Features (inc_k), feat_i \in attr (Event_j) \rightarrow Priopity (Event_j) = \max \{Severity (thr_{ks})\}, \{thr_{ks}\} \subseteq Threat (inc_k)$.

4. Метод принятия решений на основе онтологического подхода

Метод предусматривает использование онтологии, основными концептами которой являются метрики защищенности, ресурсы сети и контрмеры. Кроме типовых связей онтологии, таких как

is-a и *is-part-of*, в данную онтологию введены следующие связи: (1) между метриками и ресурсами предложены связи *has-attack* (имеется атака), *has-impact* (имеется влияние) и *has-indicator* (имеется индикатор); (2) между метриками и контрмерами предложены связи *has-impact*. Система логического рассуждения, используя онтологию метрик защищенности, позволяет сделать следующие основные выводы: (1) о нападении на основе некоторых экземпляров метрик атак; (2) об обнаружении ресурсов, на которые направлены атаки; (3) о внутренней корреляции показателей; (4) о возможных контрмерах, учитывая различные стоимостные метрики.

Тестирование онтологии метрик защищенности и основанного на ее использовании метода принятия решений осуществлялось на следующем иллюстративном примере. Предположим, что атака использует уязвимость хоста CVE-2013-002. Целью такой атаки является нарушение конфиденциальности информации, хранящейся на узле. Атака осуществляется путем сканирования порта 25 как действие *scanningTCPHost1*. Это действие является экземпляром *scanningTCP*, входящим в подкласс *Weakness* класса *AttackPattern*. Уязвимость CVE-2013-002 является экземпляром класса *Vulnerability*. Эта уязвимость вызывает переполнение буфера в компоненте Windows Forms в системы Microsoft .NET Framework, что позволяет удаленным злоумышленникам выполнять произвольный код через созданный XAML браузер приложений или созданное .NET Framework приложение, использующее неправильный подсчет объектов во время операции копирования памяти. Рассчитанное значение метрики *CVSS_Score* (подкласс «Системные метрики», класс «Метрики») для уязвимости данного хоста будет равно 9.3, что определяет общий уровень безопасности как «оранжевый». Стоимость ущерба, причиненного атакой, будет выражаться через стоимость хранения (подкласс «Стоимостные метрики», класс «Метрики») и будет оцениваться как \$10000. Для этого экземпляра атаки существует два типа контрмер: *InstallPatch* и *closePort* (подкласс «Контрмеры»). В общем случае эти контрмеры оцениваются с помощью стоимостной метрики *CountmeasureCost* (подкласс «Стоимостные метрики», класс «Метрики») как \$1000, принимая во внимание расходы на программное обеспечение. Выбор любой из контрмер реализуется на основе оценки различных метрик. На основе значений различных показателей, принимая во внимание отношение между значениями метрики хоста и стоимости контр-

мер, делается выбор контрмеры *InstallPatch*, которая изменяет общий уровень безопасности на «зеленый». Таким образом, данный пример показывает возможность поддержки решений на основе предложенной онтологии, предлагая эффективные решения для различных типов атак.

5. Архитектура реализованной системы

Представленная система методов была реализована в системе мониторинга и управления инцидентами нового поколения, архитектура которой и цикл обработки событий представлены на рис. 1.

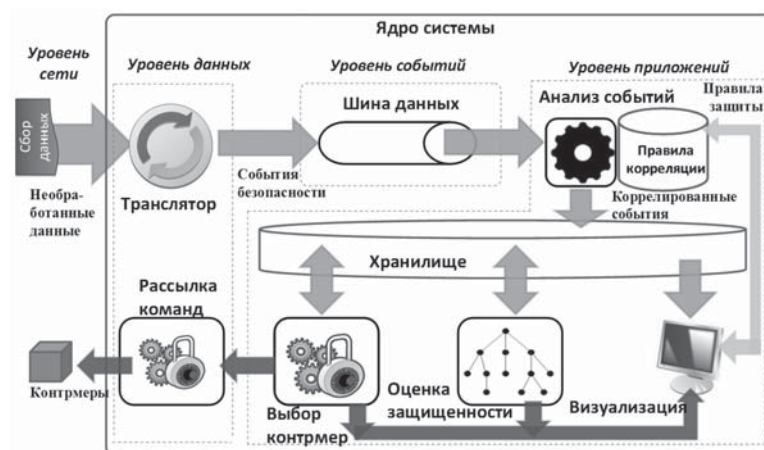


Рис. 1. Архитектура интеллектуальной системы мониторинга и управления инцидентами

В качестве тестовых областей применения в работе были рассмотрены задачи защиты таких приложений, как компьютерная инфраструктура высокой производительности, соответствующая инфраструктуре Олимпийских Игр, территориально распределенная компьютерная инфраструктура, в которой сбор и доставка данных о событиях безопасности осуществляются через телекоммуникационную среду, подвергающуюся многочисленным воздействиям, критическая информационная инфраструктура для управления гидротехническим сооружением (плотиной).

Заключение

В настоящей работе проведено формальное описание системы интеллектуальных методов мониторинга и управления инцидентами кибербезопасности, служащее основой для дальнейшего успешного развития SIEM-систем. Предлагаемое в этой области решение характеризуется тем, что в качестве системообразующего элемента выделяется репозиторий данных и знаний о событиях кибербезопасности, неотъемлемым компонентом которого является вспомогательная память (база знаний) инцидентов безопасности. Все интеллектуальные методы системы разделяются на три группы – сбора, предварительной обработки и анализа данных о событиях кибербезопасности. Для содержания методов первых двух групп приведено общее формальное описание с использованием введенных операций управления инцидентами: создания инцидента, сравнения с инцидентом, определение подобности инциденту. Представлены архитектура системы мониторинга и управления инцидентами и цикл обработки событий.

СПИСОК ЛИТЕРАТУРЫ

[Варламов, 2006] Варламов О.О. Интеллектуальные системы информационной безопасности и системный синтез модели компьютерных угроз // Искусственный интеллект». – 2006, № 3, с. 720–726.

[Галатенко, 2007] Галатенко А.В. Автоматные модели защищенных компьютерных систем // Интеллектуальные системы. – 2007. Т. 11. Вып. 1–4. С. 403–418.

[Котенко и др., 2012] Котенко И.В., Саенко И.Б., Полубелова О.В., Чечулин А.А. Применение технологии управления информацией и событиями безопасности для защиты информации в критически важных инфраструктурах // Труды СПИИРАН. – 2012. № 1. С. 27–56.

[Bridgewater, 2004] Bridgewater D. Standardize messages with the Common Base Event model, IBM. – <http://www.ibm.com/developerworks/autonomic/library/ac-cbe1/index.html>.

[CCTA, 2000] CCTA. Best Practice for Service Support. London: The Stationery Office, 2000. 306 p.

[Common, 2008] Common Event Expression White Paper. June 2008. The MITRE Corporation. 30 p.

[Miller et al., 2011] Miller D.R., Harris S., Harper A.A., VanDyke S., Blask C. Security information and event management (SIEM) implementation, 2011.

[Open, 2011] Open Group, the Distributed Audit Services. – <http://www.opengroup.org/projects/security/xdas>.

УДК 004.89

ОНТОЛОГИЧЕСКИ-ОРИЕНТИРОВАННАЯ СИСТЕМА ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ТЕХНИЧЕСКИХ ДОКУМЕНТОВ¹

Р.А. Субхангулов (*subkhangulov-ruslan@yandex.ru*)

А.А. Филиппов (*al.filippov@ulstu.ru*)

Ульяновский государственный технический университет, Ульяновск

В статье представлено описание программной системы, которая позволяет выполнять структуризацию и информационный поиск технической текстовой документации, основываясь на текущем состоянии предметной области в виде прикладной онтологии. Показана структура онтологии, перечислены основные функции системы. Приведены результаты экспериментальных исследований, доказывающие эффективность разработанной программной системы.

Введение

Повышение конкурентоспособности современных изделий, выпускаемых предприятиями, невозможно без сокращения сроков выполнения научно-исследовательских и опытно-конструкторских работ. В работах В.Н. Малюх и И.П. Норенкова отмечается, что именно на начальных этапах разработки новых систем принципиально важным является использование опыта предыдущих проектов, зафиксированных в технических документах. Решение указанной проблемы может основываться на применении интеллектуальных методов и алгоритмов анализа технических документов проектной организации с целью организации функций информационного поиска и построения навигационной структуры электронного архива.

Для эффективного применения методов интеллектуального анализа текстовой конструкторской и технической документации недостаточно рассматривать отдельный документ как набор терминов из ограниченной предметной области. Количество используемых терминов в документах настолько велико, что применение

¹ Работа выполнена при финансовой поддержке РФФИ (проекты № 14-01-31086 мол_а и № 14-01-31080 мол_а).

известных методов кластеризации и информационного поиска текстовых документов является затруднительным по причине их невысокой эффективности [Наместников, 2009]. При решении подобного рода задач применяются интеллектуальные системы, функционирование которых основано на предметно-ориентированных знаниях. Эти знания могут быть представлены в виде онтологии предметной области [Добров и др., 2006].

Учет специфики проектных знаний приводит к необходимости формирования онтологии электронного архива особой структуры, включающей в себя систему понятий предметной области, семантические отношения между ними и функции интерпретации. Поскольку любая проектируемая система изменяет свое состояние в соответствии с жизненным циклом, электронный архив должен обладать функциями адаптации к различным этапам (стадиям) жизненного цикла.

1. Модель онтологии

Формально онтология электронного архива состоит из двух прикладных онтологий и записывается как кортеж вида:

$$O = \langle O^D, O^{LC}, R_A \rangle, \quad (1)$$

где O^D – компонент онтологии предметной области, O^{LC} – онтология жизненного цикла проектируемых систем, R_A – отношение однонаправленной ассоциации между компонентами онтологии. Рассмотрим более подробно компоненты онтологии электронного архива (1).

Онтологию предметной области запишем в виде кортежа:

$$O^D = \langle C, W, R^D, F^D \rangle,$$

где C – множество понятий электронного архива, которое образует основу понятийного аппарата проектирования сложной системы; $W = W^S \cup W^P$ – множество терминов предметной области (W^S – множество терминов на уровне стандартов, W^P – множество терминов на уровне проектов); R^D – множество отношений:

$$R^D = \{R_G^D, R_C^D, R_A^D\},$$

где R_G^D – антисимметричное, транзитивное, нерефлексивное бинарное отношение обобщения («subclass_of»); R_C^D – бинарное

транзитивное отношение композиции («part_of»); R_A^D – бинарное отношение однонаправленной ассоциации.

Онтология жизненного цикла как компонента кортежа (1) записывается следующим образом:

$$O^{LC} = \langle M^{LC}, St^{LC}, R^{LC} \rangle,$$

где M^{LC} – множество моделей жизненного цикла проектируемых систем; St^{LC} – множество стадий (этапов) жизненного цикла. Отношение R^{LC} имеет вид «часть-целое (part_of)» и позволяет декомпозировать стадии жизненного цикла проектируемой системы в онтологии на этапы и т. д.

2. Функциональные возможности системы

В основе онтологического индексирования технического документа лежит следующая функция:

$$F_{oV^d} : s_i^d \rightarrow oV_{s_i^d}^d,$$

где s_i^d – i -й раздел технического документа d , $oV_{s_i^d}^d$ – онтологическое представление i -ого раздела технического документа d .

Под степенью выраженности понятия онтологии электронного архива понимается степень совпадения терминологического окружения понятия с набором терминов некоторого фрагмента технического документа при условии, что в терминологическое окружение включены термины, наиболее близкие в семантическом отношении с понятием.

Вычисление степеней выраженности понятий онтологии для каждого раздела технического документа производится с применением аппарата нечетких соответствий [Берштейн и др., 2005]. Образом множества \tilde{W}^d (множество терминов технического документа d) при соответствии \tilde{F} будем называть нечеткое множество $\tilde{F}(\tilde{W}^d)$ в $C^{P(S)}$, определяемое выражением:

$$\tilde{F}(\tilde{W}^d) = \{ \langle \mu_{\Gamma(w^d)}(c^{P(S)}), c^{P(S)} \rangle \mid c^{P(S)} \in C^{P(S)} \}, \quad (2)$$

где $\mu_{\Gamma(w^d)}(c) = \bigvee_{w^d \in W^d} (\mu_{w^d}(w^d) \& \mu_{O}(w, c^{P(S)}))$ [Берштейн и др., 2005].

Для нахождения формальной меры расстояния между документами представим каждое онтологическое представление документа

в качестве дерева (иерархии) понятий предметной области. Такая иерархия определяется путем нахождения минимального дерева, включающего все понятия из онтологического представления [Загоруйко, 1999].

Редакционное расстояние между иерархиями определяется на основе вычисления стоимости редакционной операции, которая находится отдельно для каждого типа семантического отношения. Итоговое редакционное расстояние между иерархиями вычисляется по следующей формуле:

$$\tau_{oV}^* = \max_i \left(\sum_{s=1}^m \phi_{S_i} (R_G^D)_s + \sum_{l=1}^n \phi_{S_i} (R_C^D)_l \right),$$

где $\phi_{S_i} (R_G^D)$ – стоимость редакционной операции отношения обобщения, $\phi_{S_i} (R_C^D)$ – стоимость редакционной операции отношения «часть-целое», i – номер группы стандартов; s – номер добавляемого отношения обобщения; l – номер добавляемого отношения «часть-целое», S_i – принадлежность значения редакционной операции к i -й группе стандартов.

Коэффициент нормализации T_{oV} рассчитываем исходя из всех семантических отношений обобщенной иерархии. Мера расстояния между онтологическими представлениями технических документов определяется с помощью следующего выражения:

$$\|oV^{d1} - oV^{d2}\| = \frac{\tau_{oV}^*}{T_{oV}}.$$

Для выполнения процесса формирования навигационной структуры в виде вложенного набора кластеров технических документов необходимо решить задачу настройки весов семантических отношений между понятиями онтологии на уровне стандартов, при которых качество кластеризации, определяемое выражением:

$$F^* = \frac{\max(\bar{K}_+ + \bar{K}_-, \hat{K}_+ + \hat{K}_-)}{N} \rightarrow \min, \quad (3)$$

было бы наилучшим. В выражении (3) \hat{K}_- и \bar{K}_- – множества отсутствующих документов соответственно в первом и во втором кластерах, \hat{K}_+ и \bar{K}_+ – множества лишних документов соответственно в первом и во втором кластерах, N – количество документов.

Центральное место в информационном поиске занимает понятие запроса. Запрос – это формализованный способ выражения

информационных потребностей пользователем системы. Запрос будем представлять в виде множества $Q = \{w_i, \mu_i\}$, где w_i - термины из словаря предметной области W ; μ_i - величина, характеризующая степень выраженности термина в понятии c_i из состава онтологии.

Документы электронного архива представляются в виде матрицы M_{DC} :

$$M_{DC} = \begin{matrix} & c_1 & c_2 & \cdots & c_n \\ \begin{matrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{matrix} & \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{k1} & \mu_{k2} & \cdots & \mu_{kn} \end{bmatrix} \end{matrix},$$

где d_i - документ электронного архива $D = \{d_1, d_2, \dots, d_k\}$; k - количество документов в архиве; c_i - понятие онтологии; n - количество концептов в онтологии; μ_{ij} - величина, характеризующая степень выраженности понятия в документе.

В основе онтологически-ориентированного поиска лежит метод нахождения меры включения вектора понятий запроса в вектор понятий документа. Для решения данной задачи предлагается воспользоваться формулой нечетких включений множеств [Берштейн и др., 2005]:

$$\gamma(\overline{I_q}, \overline{I_d}) = \&(\mu_{I_q}(C) \rightarrow \mu_{I_d}(C)), \quad (4)$$

где $\overline{I_q}, \overline{I_d}$ - индексы запроса и документа; C - множество понятий онтологии, где $0 \leq \gamma(\overline{I_q}, \overline{I_d}) \leq 1$.

В настоящее время реализовано три варианта алгоритмов информационного поиска.

1. Онтологически-ориентированный алгоритм:

Для каждого понятия с помощью выражения (2) формируется индекс запроса M_{QC} . Вычисляется мера включения запроса Q в документ D (выражение (4)). Документы ранжируются по убыванию значения $\gamma(\overline{I_q}, \overline{I_d})$.

2. Иерархический онтологически-ориентированный алгоритм:

Отличается от предыдущего алгоритма методом расчета значения включения вектора понятий запроса в вектор понятий документа (выражение (5)).

$$\gamma(\overline{I_q}, \overline{I_d}) = \&(F(\mu_{I_q}(c)) \rightarrow \mu_{I_d}(c)), \quad (5)$$

где $F(\mu_{I_q}(c)) \rightarrow \mu_{I_q}(c)'$ – функция перехода от общего к частному, которая выполняет поиск понятий онтологии сверху вниз для нахождения значимых понятий, не имеющих потомков.

3. С учетом терминологического окружения понятия онтологии:

Вводимый пользователем запрос $Q = \{w_i, \mu_i\}$ предварительно обрабатывается системой, для уточнения в интерактивном режиме терминами из состава терминологического окружения понятий онтологии предметной области.

3. Результаты экспериментов

Для выполнения анализа результатов работы онтологически ориентированной подсистемы структуризации экспертом была подготовлена выборка, состоящая из 5017 технических документов и сгруппированная по двум основаниям классификации:

- по виду документации – 52 группы (ГОСТ 2.601, 2.602, 2.102, 2.701 и 3.1201);
- по тематике работ – 28 групп (изделия, рассматриваемые в документах).

Был построен индекс, содержащий в своем составе онтологические и традиционные представления технических документов (множество пар «термин-частота» [Маннинг и др., 2011]). Результаты экспериментов представлены на рисунке 2.

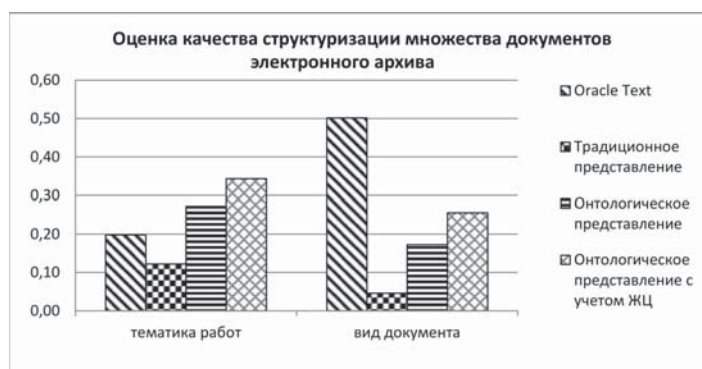


Рис. 2. Оценка качества структуризации множества документов электронного архива

Расчет качества структуризации производится с помощью выражения:

$$\hat{f}_i = 1 - \frac{\max(\sum_{i=1}^M \bar{K}_i, \sum_{i=1}^M \hat{K}_i)}{N},$$

где \bar{K}_i – множество отсутствующих документов согласно сопоставлению результатов экспертного и автоматического разбиений, \hat{K}_i – множество «лишних» документов согласно сопоставлению результатов экспертного и автоматического разбиений, $i = 1, M$ – номер кластера, M – количество кластеров, N – количество документов. Качество структуризации имеет значение от 0 до 1, где 0 – полное несоответствие экспертного и автоматического разбиений, а 1 – полное соответствие экспертного и автоматического разбиений.

Как видно из рисунка 2, лучшие значения оценочной функции для онтологических представлений с учетом моделей жизненного цикла проектируемой системы получены при структуризации выборки технических документов по тематике работ, так как данный вид структуризации электронного архива производится по содержанию отдельных документов. Для структуризации по виду документа лучшие результаты показала система Oracle Text.

Оценка качества информационного поиска определяется с помощью двух основных характеристик: *точность* поиска и его *полнота* [Маннинг и др., 2011].

Данные критерии обычно конфликтуют, стопроцентная точность и полнота на практике невозможны. Для соблюдения баланса между двумя данными параметрами применяется F -мера (F measure) [Маннинг и др., 2011]. F -мера представляет собой среднее гармоническое взвешенное (выражение (6)).

$$F = \frac{1}{\alpha \frac{1}{P} + (1 - \alpha) \frac{1}{R}} = \frac{(\beta^2 + 1) PR}{\beta^2 P + R}, \quad (6)$$

где $\beta^2 = \frac{1 - \alpha}{\alpha}$, $\alpha \in [0, 1]$, т.е. $\beta^2 \in [0, \infty]$, P – точность поиска, R – полнота поиска. При $\beta < 1$ предпочтение отдается точности поиска, при $\beta > 1$ – полноте поиска. При $\beta = 1$ выражение (6) примет вид *сбалансированной* F_1 -меры:

$$F_1 = \frac{2PR}{P+R}. \quad (7)$$

Результаты работы онтологически-ориентированной подсистемы информационного поиска (ООПИП) сравнивались со следующими поисковыми системами:

1. Яндекс. Персональный поиск (ЯПП).
2. Архивариус 3000 (А300).
3. AOL Desktop Search (AOL).
4. Copernic Desktop Search (CDS).
5. Тестовая поисковая система, основанная на векторной модели поиска (ВМ).

Для выполнения анализа результатов работы ООПИП использовалась обучающая выборка, состоящая из 41 текстового документа, 20 из которых являются техническими заданиями, а 21 – статьями из области построения информационных систем. Для расчета качества информационного поиска использовалось выражение (7). Результаты экспериментов представлены в таблице 1.

Таблица 1

Оценка качества информационного поиска множества документов электронного архива (для первых 10 документов)

Оценка	ООПИП	ВМ	ЯПП	А300	AOL	CDS
Запрос «вариантов использования»						
R	0,6	0,1	0,3	0,2	0,2	0,2
P	0,8	0,1	0,4	0,3	0,3	0,3
F-мера	0,7	0,1	0,3	0,2	0,2	0,2
Запрос «вставка оператор обновление»						
R	0,4	0,4	0,2	0,2	0	0
P	0,6	0,6	0,1	0,1	0	0
F-мера	0,5	0,5	0,2	0,2	0	0
Запрос «абстрактная визуализация объекта»						
R	0,6	0,4	1	0,8	0	0
P	0,8	0,6	0,6	0,6	0	0
F-мера	0,7	0,5	0,7	0,7	0	0

Как видно из результатов экспериментов (таблица 1), онтологически-ориентированные методы информационного поиска показывают лучшие или аналогичные результаты.

Таким образом, использование онтологически-ориентированных методов индексирования, структуризации и информационного поиска, учитывающих применяемые в процессе проектирования серии стандартов и модели жизненного цикла, предпочтительно при необходимости более качественного нахождения требуемых документов по их содержанию.

СПИСОК ЛИТЕРАТУРЫ

[**Берштейн и др., 2005**] Берштейн Л.С., Боженюк А.В. Нечеткие графы и гиперграфы. – М.: Научный мир, 2005.

[**Добров и др., 2006**] Добров Б.В., Лукашевич Н.В. Лингвистическая онтология по естественным наукам и технологиям: основные принципы разработки и текущее состояние // Десятая национальная конференция по искусственному интеллекту с международным участием (Обнинск, 25–28 сентября 2006 г.) – М.: Физматлит, 2006.

[**Загоруйко, 1999**] Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: ИМ СО РАН, 1999.

[**Маннинг и др., 2011**] Маннинг К., Рагхаван П., Шютце Х. Введение в информационный поиск. – М.: Вильямс, 2011.

[**Наместников, 2009**] Интеллектуальные проектные репозитории. – Ульяновск: УлГТУ, 2009. Наместников А.М. Интеллектуальные проектные репозитории. – Ульяновск: УлГТУ, 2009.

УДК 004.422.834

АВТОМАТИЧЕСКАЯ ГЕНЕРАЦИЯ ЗАДАНИЙ ДЛЯ КРАУДСОРСИНГОВОГО ТЕЗАУРУСА¹

Д.А. Усталов (*dau@imm.uran.ru*)
ИММ УрО РАН, Екатеринбург

Автоматическая генерация заданий для пользователей является важной проблемой коллективно создаваемых информационных ресурсов, в работе над которыми принимает участие большое количество волонтеров. В статье проанализирована данная проблема и представлен подход к ее решению на примере проекта по созданию открытого электронного тезауруса русского языка.

Введение

Задача автоматической генерации заданий состоит в построении отношения частичного порядка на множестве доступных заданий A для целевого пользователя $u \in U$ таким образом, чтобы минимизировать штраф $\sum_{u \in U} Q(A|u)$. Решение такой задачи чрезвычайно важно при выполнении широкомасштабного краудсорсингового проекта, где требуется правильно распределить работу между волонтерами и дать людям возможность работать над изолированными фрагментами поставленной работы, минимизируя вероятность возникновения конфликтов между участниками проекта.

Проект Yet Another RussNet посвящен созданию открытого электронного тезауруса русского языка на основе краудсорсинга. Основная особенность проекта заключается в использовании «сырых» данных – определений слов и списков синонимов, загруженных из доступных семантических словарей. Задача пользователей состоит в построении синсетов (множеств квазисинонимов) на основе таких «сырых» данных. По итогам эксперимента с участием студентов-лингвистов Уральского федерального университета, предполагается, что такой подход позволяет заметно ускорить процесс создания тезауруса по сравнению с традиционным экспертным подходом [Braslavski et al., 2014].

¹ Работа поддержана проектом РГНФ №13-04-12020 «Новый открытый электронный тезаурус русского языка».

1. Аналогичные работы

К сожалению, проблематика автоматической генерации заданий достаточно слабо освещена как в русскоязычной, так и в зарубежной литературе. Часто вопрос генерации заданий рассматривается в контексте подготовки заданий для учащихся школ и студентов высших учебных заведений с целью борьбы со списыванием и обеспечения индивидуализации задач в контрольных работах. Известные подходы предполагают подготовку специальных шаблонов задач, параметризуемых значениями, просчитанными заранее по заданному алгоритму [Кручинин и др., 2004; Посов, 2007]. Обозначенные подходы не подходят для задач краудсорсинга, поскольку правильное решение задачи в таком случае не может быть известно *a priori*.

Известен достаточно успешный опыт по автоматической выдаче заданий в проекте OpenCogroa, где задания по лингвистической разметке генерируются на основе текстов открытого корпуса, группируются в «пулы заданий», и пользователям выдаются задания из этих «пулов» [Бочаров и др., 2013]. Этот подход хорошо применим для жанра краудсорсинга «механизированный труд», но не для жанра «мудрость толпы», используемый в настоящее время в Yet Another RussNet (подробнее о жанровой классификации см. [Wang et al., 2013]).

Автоматическая генерация заданий для пользователей встречается и в различных платформах массовых открытых онлайн-курсов, поддерживающих режим *взаимной оценки*², где каждый пользователь должен выполнить какое-либо творческое задание и затем анонимно оценить выполнение этого задания несколькими другими участниками [Coursera, 2014]. Задача считается принятой тогда и только тогда, когда участник выполнил оба этапа работы. Детали реализации соответствующих алгоритмов не публикуются. Предполагается, что для каждого пользователя список заданий генерируется случайным образом. Этот способ может быть эффективен для эквивалентно важных заданий, но вне образовательных задач такое встречается достаточно редко.

² В англоязычных источниках этот режим называется peer assessment.

2. Генерация заданий в краудсорсинговом тезаурусе

В проекте Yet Another RussNet пользователь самостоятельно выбирает слово-задание для построения синсетов из представленного списка слов, упорядоченного по относительной частоте $freq(w)$ встречаемости слова w (рис. 1а). Такой способ генерации заданий не учитывает степень проработанности слов, наличие дополнительных «сырых» данных для обработки и т. д. Пользователи имеют склонность выбирать первые слова из списка. Это приводит к тому, что обработаны оказываются только наиболее частотные слова.

Таким образом, необходимо реализовать подход к генерации заданий, который предпочитает низкочастотные слова с большим количеством «сырых» данных высокочастотным словам с меньшим количеством имеющихся для них «сырых» данных (рис. 1а, 1б и 1в) и, по возможности, ставит менее проработанные слова над более проработанными.

Формально задача автоматической генерации заданий может быть рассмотрена как эквивалент задачи обучения ранжированию, состоящей в автоматическом подборе ранжирующей модели по обучающей выборке по отношению к каждому отдельно взятому пользователю.

Для оценивания слова w предлагается следующая формула ранжирования с учетом проработанности слова и его редакционного потенциала, состоящего в доступности «сырых» данных (1):

$$score(w) = (1 + freq(w)) \times \frac{1 + |RS(w)|}{1 + \sum_{s \in S(w) \wedge |s| > 1} |s|} \times \frac{1 + |RD(w)|}{1 + |D(w)|}, \quad (1)$$

где $freq(w)$ – частота слова w , $RS(w)$ – множество «сырых» синонимов слова w , $S(w)$ – количество синсетов со словом w , $|s|$ – количество слов в синсете s , $RD(w)$ – множество «сырых» определений слова w , $D(w)$ – множество определений слова w во всех синсетах с этим словом. Для сглаживания слишком больших значений $freq(w)$ использован десятичный логарифм, дополнительные члены-единицы добавлены во избежание неопределенного значения логарифма и деления на ноль в случае отсутствия тех или иных данных. Слова, для которых неизвестна частота, исключаются из рассмотрения.

В настоящее время в проекте Yet Another RussNet ведется работа только с именами существительными. Получено значение функ-

ции для двадцати тысяч слов, для которых известна относительная частота, полученная на основе Национального корпуса русского языка [Ляшевская и др., 2014]. Эти слова доступны для участников проекта: <http://russianword.net/editor>.

На рис. 1г представлен результат работы предложенной функции ранжирования. Из рисунка видно, что распределение слов получило более равномерный характер. Теперь низкочастотные слова, для которых существует достаточное количество «сырых» данных, в ряде случаев получили приоритет над высокочастотными словами без «сырых» данных.

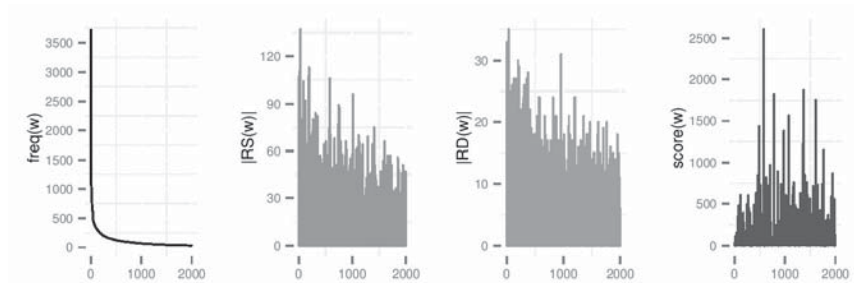


Рис. 1 Слева направо: а) первые две тысячи наиболее частотных слов по убыванию частоты, б) количество «сырых» синонимов и в) «сырых» определений для этих слов, г) значение функции ранжирования для этих слов

По сравнению с ранее использованной функцией $score(w) = freq(w)$ наблюдается заметное изменение списка слов. В частности, из вершины списка исчезли слова, для которых не имеется «сырых» данных или таких данных мало. Результат работы сбалансированной функции $score(w)$ достаточно интуитивен, однако в будущем потребуется дополнительное изучение области ее применения.

3. Особенности реализации

Функция ранжирования, представленная в формуле (1), реализована в качестве хранимой процедуры СУБД PostgreSQL на языке SQL под именем `yarn_word_score`. В связи с большим количеством разнородных данных, обрабатываемых при исполнении данной процедуры, план выполнения запроса оказался вычислительно сложным и обеспечить эффективную работу функции ранжирования в режиме реального времени не представляется возможным.

Реализация функции ранжирования в указанном варианте доступна в репозитории³ проекта YARN на GitHub в виде кода на языке программирования Ruby с использованием фреймворка для построения web-приложений Ruby on Rails.

Для предотвращения отказа в обслуживании при использовании процедуры *yarn_word_score* был использован встроенный в PostgreSQL механизм материализованных представлений, позволяющий сохранить результат выполнения запроса, построить индексы по полученным данным и работать с сохраненными результатами ранжирования как с обычной реляционной таблицей вместо хранимой процедуры, упрощая работу с данными из объектно-реляционного отображения ActiveRecord.

При создании материализованного представления *current_words_scores* каждое слово в тезаурусе получает оценку при помощи процедуры *yarn_word_score* и временную отметку о моменте вычисления функции. Для ускорения запросов к *current_words_scores* строятся соответствующие индексы по идентификатору слова и по значению функции (листинг 1).

```
CREATE MATERIALIZED VIEW current_words_scores AS
SELECT
    current_words.id AS word_id,
    yarn_word_score(current_words.id) AS score,
    LOCALTIMESTAMP AS created_at
FROM current_words;
CREATE INDEX index_current_words_scores_on_word_id
    ON current_words_scores(word_id);
CREATE INDEX index_current_words_scores_score
    ON current_words_scores(score);
```

Листинг 1. Создание материализованного представления *current_words_scores*

Обновление материализованного представления *current_words_scores* выполняется ежесуточно в полночь по Екатеринбургскому времени (UTC+6) путем полного пересчета значения *yarn_word_score* для каждого слова. Благодаря этому при активной работе список заданий меняется ежесуточно на несколько позиций.

³ https://github.com/russianwordnet/yarn/blob/develop/db/migrate/20140308194904_add_definitions_to_yarn_word_score.rb

Заключение

Представленный сбалансированный метод автоматической генерации заданий реализован, внедрен в инструмент коллективного редактирования синсетов проекта Yet Another RussNet и используется вместо первоначальной функции ранжирования на основе частоты $freq(w)$.

Дальнейшие исследования могут быть посвящены построению взвешенной сбалансированной модели и добавлению новых параметров, таких как индивидуальные и поведенческие факторы участника краудсорсингового проекта.

Благодарности. Автор благодарит анонимных рецензентов за ценные замечания по содержанию работы, а также всех участников проекта Yet Another RussNet (<http://russianword.net/>), принявших участие в работе над словами.

СПИСОК ЛИТЕРАТУРЫ

[Бочаров и др., 2013] Бочаров В., Алексеева С., Грановский Д., Протопопова Е., Степанова М., Суриков А. Морфологическая разметка корпуса силами волонтеров // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 29 мая – 2 июня 2013 г.). – М.: РГГУ, 2013.

[Braslavski et al., 2014] Braslavski P., Ustalov D., Mukhin M. A Spinning Wheel for YARN: User Interface for a Crowdsourced Thesaurus // Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics, Sweden, 2014.

[Wang et al., 2013] Wang, A., Hoang, C.D.V., Kan, M.Y.: Perspectives on crowdsourcing annotations for natural language processing // Language Resources and Evaluation. 2013. № 47(1).

[Кручинин и др., 2004] Кручинин В., Морозова Ю. Модели и алгоритмы генерации задач в компьютерном тестировании // Известия Томского политехнического университета. 2004. Т. 307. № 5.

[Coursera, 2014] Coursera Help | How do peer assessments work? – <http://help.coursera.org/customer/portal/articles/1163294-how-do-peer-assessments-work>.

[Ляшевская и др., 2014] Ляшевская О., Шаров С. О частотном словаре национального корпуса русского языка // Качество. Инновации. Образование. 2014. № 1.

[Посов, 2007] Посов И. Автоматическая генерация задач // Компьютерные инструменты в образовании. 2007. № 1.

СЕМИНАР

**БЕСПИЛОТНЫЕ ТРАНСПОРТНЫЕ
СРЕДСТВА С ЭЛЕМЕНТАМИ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

УДК 004.932.2

**АНАЛИЗ ИЗОБРАЖЕНИЙ, ПОЛУЧАЕМЫХ С БОРТОВОЙ
КАМЕРЫ БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА
ДЛЯ ЕГО НАВИГАЦИИ**

Х.А. Абдулкадим (*h_a_meer@yahoo.com*)

Ульяновский государственный технический университет
Ульяновск

В статье представлены некоторые результаты исследования возможности применения различных алгоритмов оценивания параметров пространственных смещений изображений, получаемых с бортовой камеры автономного беспилотного аппарата, для решения задач навигации автономного аппарата.

Введение

В настоящее время известно большое количество алгоритмов, позволяющих выполнять оценку смещений и искажений в паре изображений [1–4]. Среди таких алгоритмов можно упомянуть корреляционно-экстремальные алгоритмы, процедуры, основанные на выделении и сравнении характерных черт изображений, алгоритмы направленного поиска и другие. Некоторые из этих алгоритмов нашли свое применение в задачах, связанных с навигационным сопровождением различных мобильных объектов. Такое сопровождение, как правило, связано с поиском на заранее заданной электронной карте местности изображения, получаемого с помощью бортовой камеры. Несмотря на ряд очевидных преимуществ применения таких алгоритмов, связанных с отсутствием необходимости во внеш-

них по отношению к позиционируемому объекту устройствах и с постоянными невозрастающими ошибками, такая идентификация сопряжена с рядом сложностей. Среди этих сложностей высокие требования к производительности управляющих вычислительных систем и существенные для практической реализации ошибки оценивания. Это объясняется большим количеством оцениваемых параметров и сложным видом минимизируемой целевой функции. В настоящей работе представлены результаты исследований, позволяющие определить условия для применения алгоритмов оценки смещения изображений для решения задач навигации автономных летательных аппаратов.

2. Псевдоградиентные алгоритмы

Рассмотрим временную последовательность изображений $x_{ij}^{(k)}$, где i, j – пространственные координаты элемента на изображении, k – номер изображения в последовательности. Выберем в данной последовательности два изображения, следующие друг за другом. Обозначим их для простоты $x_{ij}^{(1)}$ и текущее $x_{ij}^{(2)}$. Будем называть далее $x_{ij}^{(1)}$ опорным кадром, а $x_{ij}^{(2)}$ текущим. В случае, если эти кадры получены с камеры, установленной на автономном летательном аппарате, то, оценив разницу между ними, становится возможно определить относительное перемещение самого летательного аппарата за прошедший между моментами регистрации промежуток времени.

Будем далее считать, что шум, воздействующий на опорное изображение $x_{ij}^{(1)}$ и текущее $x_{ij}^{(2)}$, является белым, так что наблюдения для первого и второго изображений имеют следующий вид:

$$\begin{aligned} Z_{ij}^{(1)} &= x_{ij}^{(1)} + \theta_{ij}^{(1)}, \\ Z_{ij}^{(2)} &= x_{ij}^{(2)} + \theta_{ij}^{(2)}. \end{aligned}$$

Обычно при оценивании пространственного сдвига между z^1 и z^2 используют простой алгоритм, описанный в [2]. Сама оценка смещения в соответствии с этим алгоритмом будет иметь следующий вид:

$$\hat{h} = \frac{2 \sum_{j=1}^N (z_{ij}^{(1)} - z_{ij}^{(2)})(z_{ij+1}^{(1)} - z_{ij-1}^{(2)})}{\sum_{j=1}^N (z_{ij+1}^{(1)} - z_{ij-1}^{(2)})^2}.$$

Недостатком данной процедуры являются значительные ошибки, возникающие при смещениях большого размера. Выходом в этих условиях является применение псевдоградиентных процедур. Эти процедуры могут быть описаны следующим выражением:

$$\hat{\alpha}_{t+1} = \hat{\alpha}_t - \Lambda_{t+1} \beta_{t+1} (J(Z_{t+1}, \hat{\alpha}_t)),$$

где α – оцениваемый вектор параметров; t – номер итерации; Λ – матрица приближения; β – псевдоградиент целевой функции J , которая характеризует качество оценивания; $Z_t = \{z_{jt}^{(2)}, \hat{z}_{jt}^{(1)}\}$ – локальная выборка наблюдений, используемая на t -ой итерации; $z_{jt}^{(2)} \in Z^{(2)}$, $jt \in \Omega_j$; $\hat{z}_{jt}^{(1)} = \hat{z}_{jt}^{(1)}(j_t, \alpha_{t-1}) \in \hat{Z}$ – выборка, полученная из наблюдений $Z^{(1)}$, например, с помощью интерполяции; $\Omega_j = \{(j_x, j_y)^T\}$ – сетка наблюдений. Интервал дискретизации соответствует числу наблюдений $z_{jt}^{(2)}$.

2. Результаты исследований

Для исследований эффективности применения различных алгоритмов мы использовали несколько последовательностей изображений различных размеров, полученных с мобильных летательных аппаратов. При этом фактический объем обрабатываемых данных варьировался в зависимости от временной длительности самой видеопоследовательности и размеров каждого кадра. На рисунке 1 представлены некоторые образцы изображения для исследований.



Рис. 1. Образцы изображения

На рисунке 2 представлены зависимости дисперсии ошибки оценивания от размеров изображения и количества оцениваемых параметров (сдвиг по одной оси – нижняя кривая на графике или по двум осям – верхняя кривая на графике).

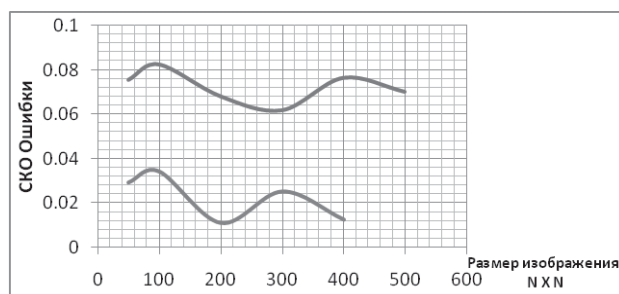


Рис. 2. Дисперсии ошибки оценивания

Анализ полученных кривых показывает, что применение псевдоградиентных алгоритмов позволяет обеспечить достаточную для большинства практических приложений точность (менее 0.3 пикселей в большинстве случаев). При этом данная точность существенно не зависит от геометрических размеров самого изображения.

На рисунке 3 представлена зависимость времени обработки от величины фактического сдвига и размера изображения. При этом для общности результатов выполнялась оценка четырех параметров – пространственного сдвига по двум осям, угла поворота и масштаба.

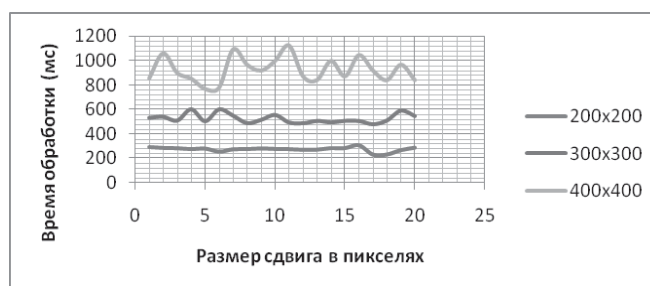


Рис. 3. Время обработки вычисляется для разных размеров

Представленные результаты свидетельствуют о линейной зависимости времени обработки от геометрических размеров кадров. При этом время обработки существенно не зависит от размера самого сдвига. Анализ самого времени обработки (менее 1 секунды) свидетельствует о потенциальной возможности применения данных алгоритмов в системах реального времени.

На рисунке 4 и рисунке 5 представлены зависимости точности получаемых оценок от размеров сдвига кадров относительно друг друга.

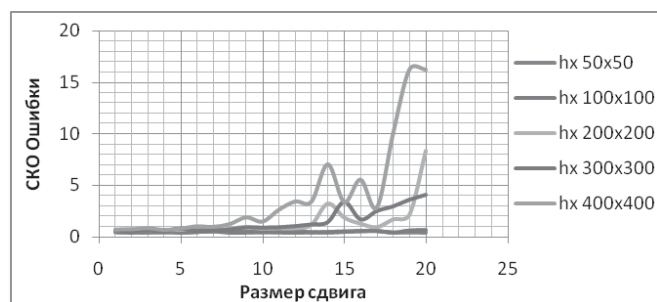


Рис. 4. Точность оценки (hx) с различными размерами изображения

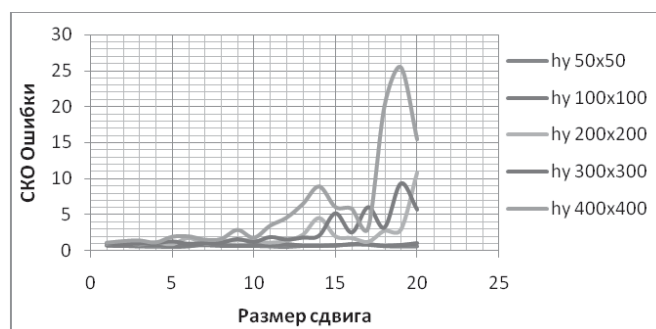


Рис. 5. Точность оценки (hy) с различными размерами изображения

Результаты свидетельствуют о резком падении точности оценивания при сдвигах, сопоставимых с радиусом корреляции. Это объясняется «сваливанием» псевдоградиента в зоны «ложных» экстремумов и существенно ограничивает применение псевдоградиентных алгоритмов. С другой стороны, это не является препятствием для оценки смещений в последовательности изображений, получаемой с бортовой камеры автономного летательного аппарата, особенно в случае достаточной высоты последнего.

Заклучение

Таким образом, в настоящей статье представлены результаты представительных статистических исследований, проведенных на реальных аэрокосмических видеопоследовательностях. Данные результаты свидетельствуют о потенциальной возможности использования быстродействующих псевдоградиентных алгоритмов для позиционирования самих аппаратов.

СПИСОК ЛИТЕРАТУРЫ

[**Kumar, 2011**] Akhilesh Kumar and Chandan Kumar, «Estimation of motion among shifted images at coarse and fine levels «International Journal of Computational Intelligence and Information Security, India, December 2011 Vol. 2, No. 12.

[**Tashlinskii, 2000**] A.G. Tashlinskii, «Image Sequence Spatial Deformation Parameter Estimation», UIGTU, ISBN 5-89146-204-4, Ulyanovsk [in Russian], 2000.

[**Кавеев, 2011**] И.Н. Кавеев. Разработка и исследование псевдоградиентных алгоритмов привязки изображений в условиях интенсивных помех. Диссертация в УЛГТУ (Ульяновский государственный технический университет), Ульяновск, 2011.

[**Horeva, 2013**] A.M. Horeva, A.G. Tashlinskii, and L. Sh. Birtimirov, «Finding the Pseudogradient of the Objective Function in Procedures for Estimation of Interframe Deformations of Images», Springer US ,ISSN 1054_6618, Pattern Recognition and Image Analysis, 2013, Vol. 23, No. 4, pp. 440–444.

УДК 681.51

**РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММНЫХ СРЕДСТВ
УПРАВЛЕНИЯ БЕСПИЛОТНЫМ ЛЕТАТЕЛЬНЫМ
АППАРАТОМ**

С.В. Воронов (*s.voronov@ulstu.ru*)

В.Е. Дементьев (*vitawed@mail.ru*)

С.С. Логинов (*ss7loginov@gmail.com*)

Ульяновский государственный технический университет
Ульяновск

В работе рассмотрено решение задачи управления беспилотным летательным аппаратом, основанное на использовании комплексирования инерциальных систем навигации и данных, поступающих с бортовой видеокамеры. Приведено описание программного комплекса управления.

В последнее время все больше внимания уделяется разработке беспилотных летательных аппаратов (БЛА). Это связано с несколькими факторами, а именно их низкой эксплуатационной стоимостью, простотой использования, а также широким спектром решаемых задач. Среди множества реализаций БЛА особый интерес представляют мультикоптеры – летательные аппараты с несколькими несущими винтами, вращающимися диагонально в противоположных направлениях. Мультикоптеры обладают высокой скоростью, маневренностью, легкостью конструкции, а также простотой в эксплуатации. Некоторые образцы способны до часа находиться в полете и поднимать вес до 10 кг. Они незаменимы для съемки с воздуха. Современные мультикоптеры обладают сложной бортовой микропроцессорной системой, сопряженной с множеством разнообразных датчиков.

Проблемой автоматического управления БЛА занимаются как в нашей стране, так и за рубежом. На данный момент большинство программных решений по автоматическому пилотированию БЛА осуществляется с помощью систем глобального позиционирования GPS/ГЛОНАСС. К недостаткам таких систем можно отнести относительно низкую точность позиционирования, значительные задержки сигналов со спутника, невозможность работы внутри помещений, отсутствие реакции на препятствия и т.д. Кро-

ме того, существенной проблемой, препятствующей появлению и распространению роботов-мультикоптеров, является отсутствие эффективных быстродействующих алгоритмов управления, позволяющих роботу перемещаться в условиях ветровых возмущений и сложного рельефа местности. Для решения этих и других задач в настоящей работе предлагается использовать механизмы комплексирования навигационных систем, разработанные нами при решении различных задач управления автономными объектами [Vasil'ev et al., 2013], [Васильев, 2011].

Разрабатываемые нами алгоритмы предполагают проведение совместной обработки данных с датчиков, установленных на борту БЛА. Среди этих датчиков ключевыми являются: бортовой гироскоп, позволяющий определять элементарные угловые смещения аппарата, магнитный компас, дающий информацию об абсолютных углах наклона, инерциальная навигационная система, оценивающая относительные смещения аппарата, датчики высоты. Получаемый с этих устройств поток данных комплексировается с результатами обработки видеопотока с бортовой видеокамеры. Обработка выполняется с помощью быстродействующих псевдоградиентных алгоритмов, позволяющих БЛА определять свое местоположение относительно наблюдаемых объектов. Координаты текущего местоположения сравниваются с координатами цели, после чего происходит корректировка курса.

При реализации данных алгоритмов нами использовался квадрокоптер Ar Drone 2.0.

Программный комплекс представляет собой набор функциональных модулей, выполненных в виде отдельных исполняемых файлов и взаимодействующих между собой посредством UDP-протокола.

Модуль планирования миссии позволяет задавать в пространстве маршрут движения БЛА, а также служит для запуска и остановки процесса моделирования всего программного комплекса. При запуске процесса моделирования сформированный маршрут движения БЛА передается по UDP-протоколу блоку контроля исполнения миссии.

У каждого блока имеются порты на передачу и прием информации, которые связаны с соответствующими портами других блоков. В результате получается множество независимых информационных

потоков между блоками, каждый из которых работает в отдельном программном потоке.

При передаче блоку контроля исполнения миссии управления движением команды на начало полетной программы выполняется начальная инициализация данного блока и запускается процесс управления движением.

Блок контроля исполнения миссии и управления движением служит для контроля и выполнения команд миссии, которые представляют собой указания для блока управления движением о требуемом изменении положения БЛА, таких как перейти к указанной точке, встать на заданный курс, остановиться и т. п. В случае возникновения препятствий на пути движения БЛА оператор имеет возможность экстренно посадить БЛА.

Для определения координат текущего положения агента в пространстве используются навигационные устройства БЛА, которые принимаются по UDP-протоколу и имеют следующую структуру:

Header 0x55667788	Drone state	Sequence number	Vision flag	Option 1			...	Checksum block		
				id	size	data		cks id	size	cks data
32-bit int.	32-bit int.	32-bit int.	32-bit int.	16-bit int.	16-bit int.	16-bit int.	16-bit int.	32-bit int.

Рис. 1. Структура данных, поступающих с БЛА

Затем в блоке комплексирования строится оценка координат положения на основе текущей сенсорной информации. Вычисленная оценка передается блоку управления движением и служит для определения достижения агентом заданной координаты.

Блок обработки данных по серии ТВ изображений определяет изменение положения БЛА по смещениям изображений в кадрах ТВ-системы, с последующей передачей информации блоку комплексирования.

Блок распознавания осуществляет идентификацию объектов на пути движения БЛА. Распознавание осуществляется на основе цветовой фильтрации поступающих изображений. При этом фильтрация может выполняться как в цветовом пространстве RGB, так и в пространстве HSV, выбор между которыми основывается на характеристиках выделяемого объекта.

Оценка изменения местоположения БЛА по последовательности кадров строится на основе оценивания межкадровых геометри-

ческих деформаций соседних кадров. В качестве алгоритма оценивания используются безыдентификационные псевдоградиентные процедуры вида [Ташлинский, 2011]:

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \beta_t (J(\hat{\alpha}_{t-1}, Z_t)),$$

где β – псевдоградиент целевой функции J ; $t = \overline{1, T}$ – номер итерации; Λ_t – матрица усиления; Z_t – локальная выборка отсчетов исследуемых изображений, используемая для нахождения β на t -й итерации. Эти алгоритмы рекуррентны, позволяют достичь высокой точности оценивания при небольшом объеме вычислительных затрат, не требуют предварительной оценки параметров исследуемых изображений, а также применимы при обработке изображений в условиях априорной неопределенности и устойчивы к различным помехам.

Вследствие высокой скорости поступления кадров вероятность появления значительных яркостных искажений на паре исследуемых изображений очень мала. Следовательно, в качестве целевой функции ввиду небольших требований к вычислительным затратам целесообразно использовать средний квадрат межкадровой разности [Taslinskii, 2007]. Существует несколько способов нахождения псевдоградиента данной целевой функции, которые подробно описаны в работе [Ташлинский и др., 2013].

Для поставленной задачи оцениваемыми параметрами для этих алгоритмов будут координаты центра объекта на изображении (вектор пространственного смещения), расстояние до объекта (точнее, проекция этого расстояния на горизонтальную плоскость) и угол поворота на объект. Оценив эти параметры и зная фактическое месторасположение эталонного объекта в пространстве, можно однозначно определить текущие координаты БЛА. Выражения для функции искажения координат и процедура нахождения параметров деформации описаны в работе [Дементьев и др., 2013].

Применение указанных процедур позволяет окончательно сформировать следующую функциональную схему совместной обработки навигационной информации (рис. 2).

При этом сама выполняемая миссия представляет собой маршрут движения БЛА и состоит из набора точек с заданными координатами, которые связаны между собой линейными сегментами. При запуске процесса моделирования блоку контроля исполнения

миссии и управления движением передается маршрут (миссия), сформированный оператором системы. Целью данного блока является реализация алгоритма по контролю прохождения БЛА переданного маршрута.

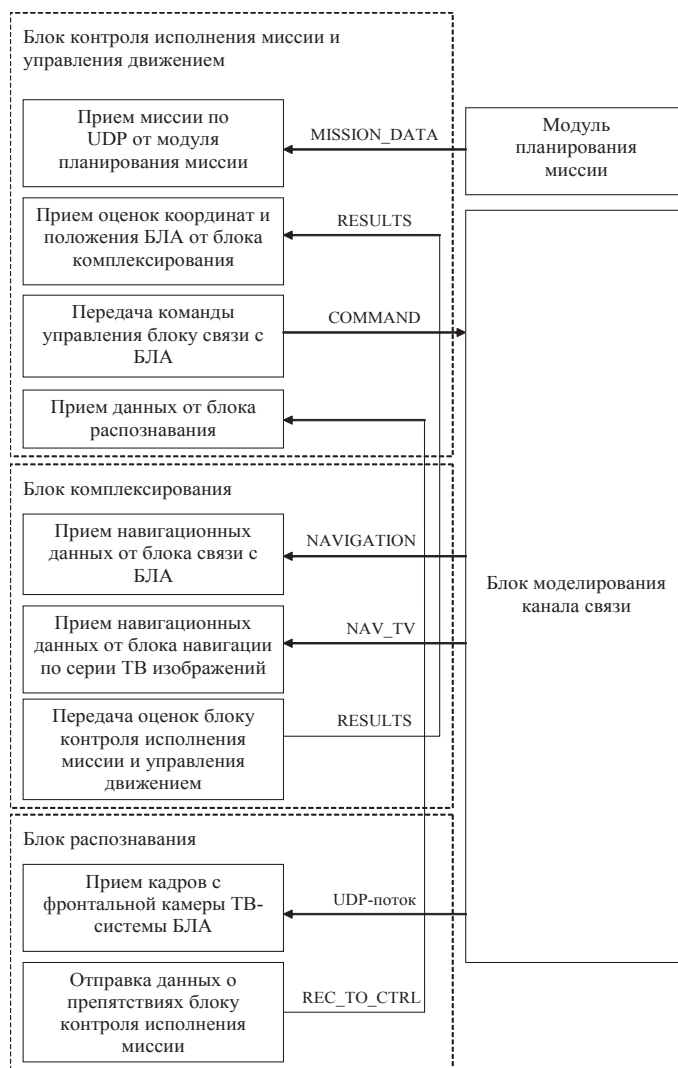


Рис. 2. Схема взаимодействия блоков управления, распознавания и комплексирования с другими блоками программного комплекса

Для реализации поставленной задачи маршрут переводится на уровень команд, передаваемых системе управления движением, представляющих собой указания:

- перейти к заданной точке с требуемой скоростью;
- изменить курс;
- остановиться;
- взлететь/приземлиться;

и т.п.

В каждый конкретный момент времени система управления движением может выполнять только одну из команд. При этом, если в момент выполнения текущей команды от контроллера исполнения миссии поступает новая команда, то система управления движением прерывает выполнение предыдущей и начинает выполнение новой команды. Благодаря такому взаимодействию блок контроля исполнения миссии имеет возможность экстренно прервать выполнение текущей команды и выполнить другие, например, экстренную посадку.

Программа для контроля БЛА работает по следующему алгоритму:

- 1) Получение координат миссии;
 - 2) Взлет.
- В цикле
- 3) Расчет направления движения до следующей точки;
 - 4) Поворот в заданном направлении;
 - 5) Расчет длины траектории;
 - 6) Движение в заданном направлении до следующей точки.
- Конец цикла
- 7) Посадка.

В реальных условиях возникает сложность реализации алгоритма из-за специфики летающих объектов, а именно наличие инерции при повороте и при остановке аппарата. Для компенсации этих процессов была предусмотрена корректировка положения аппарата.

Алгоритмы не требуют больших вычислительных мощностей, что позволяет использовать их на борту самого аппарата. Так совместная обработка данных датчиков и видео позволяет обеспечить точность позиционирования до 15 см по любой оси. Разработанные алгоритмы могут быть применены для любых моделей мультикоптеров.

СПИСОК ЛИТЕРАТУРЫ

[Vasil'ev et al., 2013] Vasil'ev K.K., Dement'ev V.E., Luchkov N.V., Analysis of Efficiency of Detecting Extended Signals on Multidimensional Grids // Pattern Recognition and Image Analysys. – 2013. – Vol.23. No. 1 – pp. 1–10.

[Васильев, 2011] Васильев К.К. Разработка алгоритмов навигации автономного аппарата, «Интегрированные автоматизированные системы управления». – Ульяновск: ФНЦП ОАО «Марс», 2011, с. 93–96.

[Ташлинский, 2011] Ташлинский А.Г. Оценивание параметров пространственных деформаций последовательностей. – Ульяновск: Издательство УлГТУ, 2000.

[Taslinskii, 2007] Taslinskii A.G. Pseudogradient Estimation of Digital Images Interframe Geometrical Deformations. // Vision Systems: Segmentation & Pattern Recognition. – Vienna, Austria: I-Tech Education and Publishing, 2007.

[Ташлинский и др., 2013] Ташлинский А.Г., Воронов С.В. Анализ целевых функций при рекуррентном оценивании межкадровых геометрических деформаций изображений // Научно-технологические проблемы радиоэлектронного приборостроения. – 2013. – Т. 14. – № 5.

[Дементьев и др., 2013] Дементьев В.Е., Воронов С.В. Применение алгоритмов цифровой обработки видеоизображений для управления автономным роботизированным аппаратом. Труды Международной научно-технической конференции «Фундаментальные проблемы радиоэлектронного приборостроения «INTERMATIC– 2013», ч. 4.

УДК 004.93'12

**РАЗРАБОТКА АЛГОРИТМОВ НАВИГАЦИИ
АВТОНОМНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ НА
ОСНОВЕ КОМПЛЕКСИРОВАНИЯ РАЗНОРОДНЫХ
НАВИГАЦИОННЫХ ДАННЫХ**

В.Е. Дементьев (*vitawed@mail.ru*)
Д.С. Кондратьев (*kondratev.dmitriy@gmail.com*)
УлГТУ, Ульяновск

В работе описывается применение алгоритмов поиска объектов на изображении, полученном с инфракрасного дальномера, а также комплексирование этой информации с бортовыми инерциальными системами в целях определения собственного местоположения.

Введение

Современные системы управления предназначены для управления сложными многофункциональными объектами, действующими в постоянно меняющейся окружающей обстановке. Это в полной мере относится и к системам управления современными автономными беспилотными летательными аппаратами (БПЛА). В состав современных БПЛА в настоящее время входит множество датчиков, потенциально способных в каждый момент времени с разной степенью достоверности определять местоположение БПЛА. Среди таких датчиков можно отметить инерциальные системы (ИНС), способные с большой точностью определять относительное смещение БПЛА, системы спутниковой навигации, видеокамеры, позволяющие получать визуальную информацию об окружающей БПЛА обстановке, ИК датчики, с помощью которых возможно определение расстояний до расположенных рядом объектов и т.д. Несмотря на разнообразие датчиков, при отсутствии сигнала от спутниковых навигационных систем аппарат теряет возможность определять собственное положение и может опираться только на автономные навигационные системы. Но до последнего времени не существовало эффективных решений, связанных с комплексированием и автоматической обработкой информации, получаемой с них. Это во многом объясняется несовершенством измерительной

аппаратуры и непреодолимыми трудностями, возникающими при обработке большого объема данных в режиме реального времени. Однако появление нового поколения датчиков и быстрой вычислительной техники привело к идее автоматизации ряда традиционных функций по обработке информации, ранее считавшихся принципиально недоступными для автоматических систем.

1. Алгоритм комплексирования навигационных данных

Одним из датчиков, потенциально применимых на БПЛА, является инфракрасный дальномер. Как было отмечено, ИК датчик используется для формирования массива расстояний до объектов. Эта информация, представляющая своеобразное полутоновое изображение, является объектом цифровой обработки, в ходе которой должно быть выполнено распознавание объектов, находящихся впереди автономного аппарата. При этом необходимо принять решение о потенциальной опасности объекта и возможности его использования в процессе определения собственного местоположения.

В качестве ИК датчика возможно использование контроллера Microsoft Kinect. Данный контроллер позволяет формировать массив расстояний до объектов с высокой точностью. В своем составе Kinect имеет инфракрасный лазерный проектор, который создает в пространстве сетку из точек и камеру, снимающую в инфракрасном спектре, которая регистрирует изображение этой сетки. Принцип формирования изображения следующий.

Излучатель(проектор) создает очень короткую равнонаправленную световую вспышку (по длительности эквивалентную расстоянию в несколько метров, что определяет максимально дальность камеры, например 7,5 метра при длительности вспышки $t_0 = 50$ нс). Отраженный от сцены свет возвращается обратно и фокусируется оптической системой так, чтобы падать на цифровую матрицу(камеру). Эта матрица имеет специальный тип и состоит из фотодиодов. При попадании света на фотодиод он начинает вырабатывать ток, который через ключ передается на накопительные элементы (конденсаторы). В простейшем случае можно использовать два конденсатора.

Согласно заявленным техническим характеристикам, контроллер Kinect поддерживает рабочий диапазон дальностей от 1 до 6 метров и имеет поле зрения примерно 50 угловых градусов в обоих направлениях. Разрешение дальномерной матрицы невелико и составляет 320 x 240 пикселей.

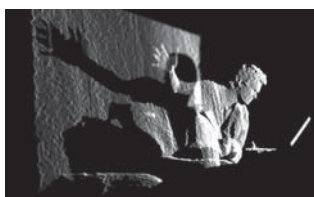


Рис. 1. Пример изображения, полученного с контроллера Microsoft Kinect

Для обработки данного изображения с целью поиска объектов и последующей навигации возможно использование следующего алгоритма, состоящего из 2-х этапов.

На первом этапе выполняется выделение в получаемых с ИК датчика эхо изображениях объектов, потенциально подходящих для использования при последующем позиционировании. Рассмотрим процедуры, выполняемые в рамках данного этапа, более подробно.

1. Считая, что датчик установлен ровно в горизонтальной плоскости (нет крена влево/вправо, а если он есть, его можно компенсировать по данным встроенных гироскопов), рассчитаем медиану каждой строчке изображения. Медиана рассчитывается следующим образом: массив данных составляющих каждую строчку сортируется, медианой считается элемент с номером $n/2$ (где n число элементов в строке – т.е. число столбцов). Таким образом, получим m медиан (где m – число строк, обычно при этом $m=n$).

2. Определим линию горизонта, то есть фактически разделим изображение на две области: «верхнюю», где объектов заведомо нет (там все «темно»), и нижнюю, потенциально содержащую объекты. Определение линии горизонта осуществляется достаточно просто. Для этого достаточно пробежаться «сверху вниз» по каждому столбцу. Как только значение яркости станет достаточно большим, то данная точка будет входить в линию горизонта (рис. 2, серая линия).

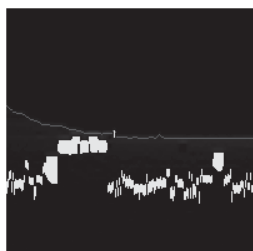
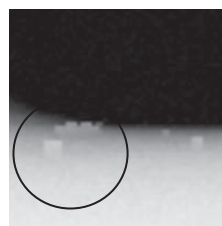


Рис. 2. Определение линии горизонта

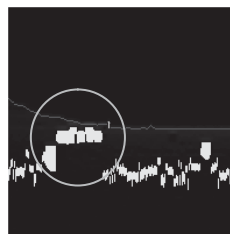
3. Определим области постоянного уровня яркости. Важность этих областей обуславливается тем, что искомые на эхо-изображениях объекты, как правило, имеют вертикальные поверхности. Эти поверхности на эхо-изображениях представляются как области с близким к постоянному яркостному заполнению.

Рис. 3. Обведены области, обладающие постоянным уровнем яркости



Для определения указанных областей выполним постолбцовый анализ от линии горизонта вниз до нижней границы изображения. Если на протяжении достаточно протяженного (больше 5 точек) участка яркость меняется незначительно (например, не более чем на 5 единиц), то считаем, что эта область является зоной потенциального расположения объекта. На следующем рисунке эта область помечена светло-серым цветом (одна из областей выделена серой окружностью).

Рис. 4. Белым цветом выделены области, являющиеся зоной потенциального расположения объектов



4. Определим зоны вертикальных границ объектов. Для этого аналогично предыдущему пункту анализируем составляющие изображения столбцы. В случае если разница между соседними (или через один) пикселями больше некоторого порога, можно считать, что выделенная точка – часть верхней границы объекта. При этом в зависимости от этой разницы и величины, на которую текущая точка превосходит медиану строки, эта точка, содержащая область возможного расположения объекта, может быть разной. Так, если

разница между точками в столбце невелика, то можно ожидать малый объект под этой точкой, если разница значительна, то объект большой. Если точка на потенциальной границе превосходит по яркости точки на текущей строке (больше, чем медиана строки), то это также показатель размера объекта. В ходе большого количества экспериментов была получена следующая оценка потенциального размера объекта: $(d \div 8) + (m \div 8)$, где d – вертикальная разница на границе потенциального объекта, m – разница между значением яркости на границе и медианы строки. На следующем рисунке показаны области возможного расположения объектов.

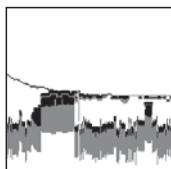


Рис. 5. Определение зон вертикальных границ

5. Определяем горизонтальные границы объекта аналогично п. 4. При этом для каждого объекта определяем и вертикальную, и горизонтальную границы объекта. На следующем рисунке показаны области горизонтальных границ объекта. При этом такой границ может быть затруднен малыми значениями разностей яркости «у основания объекта» (потому что у своей «нижней границы» объект для эхолота практически растворяется на фоне поверхности). В связи с этим на рисунке правая граница объекта в некоторых местах не найдена (объект ложно растянут до правого края изображения).

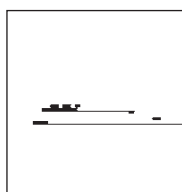
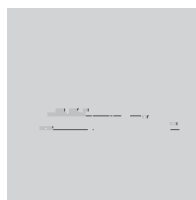


Рис. 6. Определение горизонтальных границ объекта

6. Определяем пересечения найденных в предыдущих трех пунктах областей путем проверки простейших условий (если точка попала и в область 1, и в область 2, и в область 3, то считаем, что эта точка принадлежит объекту).

Рис. 7. Результат пересечения областей



7. Исключаем отдельные точки в областях, не имеющих вертикальных и горизонтальных «соседей» (фильтруем выбросы). На предыдущем рисунке такие точки показаны черным цветом.

8. Определяем центры полученных областей. Для этого последовательно бежим по точкам изображения, если «натыкаемся» на точку, принадлежащую области сигнала, рекуррентно определяем соседей этой точки, и их вертикальные и горизонтальные координаты записываем в специальный массив чисел. Как только распространяться «заливке» становится больше некуда, рассчитаем средние по этим двум массивам. Эти средние и будут координатами формируемого точками объекта. На рисунке они показаны белыми точками. Это и есть центры тяжести объектов.

Вторым этапом обработки получаемых данных является собственно позиционирование аппарата относительно выделенных объектов. Предположим при этом, что имеется информационный массив, содержащий характеристики объектов, обнаруженных алгоритмом обнаружения объектов на эхо-изображениях.

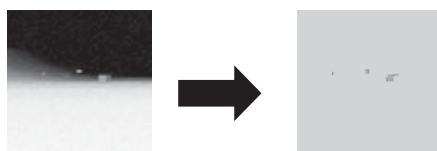


Рис. 8. Обнаруженные объекты на гидроакустическом изображении

Каждому найденному объекту на изображении соответствует расстояние от БПЛА до центра этого объекта. Зная это расстояние и смещение центра объекта относительно центра эхо-изображения, становится возможным определение расстояния до плоскости, которая проходит через данный центр объекта и расположена перпендикулярно лучу, сканирующему окружающую обстановку в центре гидроакустического изображения. При этом мы считаем, что

БПЛА всегда имеет тангаж 0° и каждый пиксел гидроакустического изображения соответствует смещению луча сканирования на 1° .

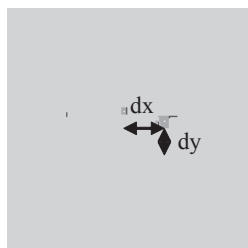


Рис. 9. Определение смещения объекта относительно центра главного (центрального) луча эхолота

Определим угол между лучом, направленным к объекту, и лучом, находящимся в центре гидроакустического изображения,

$$dh = \sqrt{dx^2 + dy^2},$$

и расстояние до плоскости, содержащей данный объект

$$R = r * \cos(dh),$$

где r – расстояние до объекта.

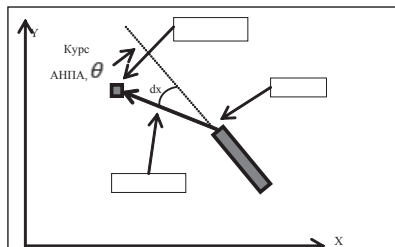


Рис. 10. Вычисление глобальных координат объекта

Считывая глобальные координаты и курс БПЛА с его инерциальных систем, определим приближенно глобальные координаты найденных объектов. На основе курса БПЛА(Θ) и смещения объекта относительно центра гидроакустического изображения в горизонтальной плоскости (угол dx) найдем направление луча от эхолота к объекту в глобальных координатах.

$$dR = dx + \Theta,$$

где dR – направление луча к объекту в глобальных координатах.

Вычислим глобальные координаты объекта

$$X = R * \cos dR \quad Y = R * \sin dR.$$

Заключение

Таким образом, имея последовательность эхо-изображений совместно с информационным массивом относительно выделенных на них объектов, можно организовать траекторное слежение за объектами, определять координаты этих объектов с заданной точностью, оценить его пригодность для использования в навигации, а также определить, является ли он неподвижным. В случае успешной идентификации объекта, пригодного для позиционирования (т.е. неподвижного и обнаруживаемого в последовательности эхо-изображений), система сможет скорректировать свои координаты по координатам объекта, полученным на предыдущих кадрах. На рис. 11 представлена реализация такого подхода, выполненная в рамках разработанного в Ульяновском государственном техническом университете комплекса моделирования и управления автономными объектами.

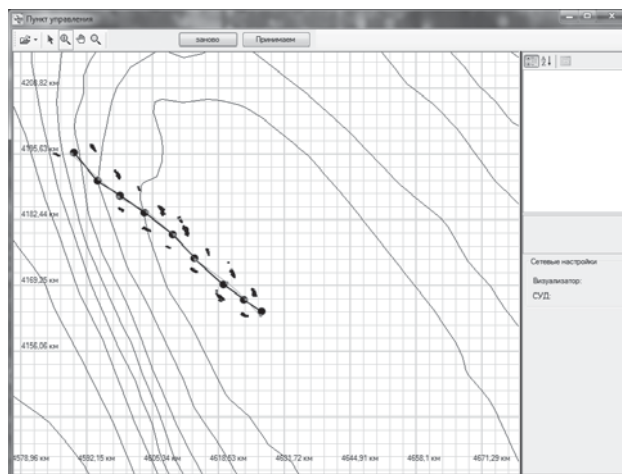


Рис. 11. Реализация алгоритма в виде программного модуля (черным цветом отмечен истинный путь аппарата, серым – координаты и путь, определенные представленным алгоритмом, точки около маршрута – обнаруженные объекты)

Впоследствии БПЛА может составить детализированную карту воздушного пространства, и с каждым разом добавлять в нее новые корректировки.

СПИСОК ЛИТЕРАТУРЫ

[Васильев, 1995] Васильев К.К. Методы обработки сигналов и полей // Межвузовский сборник научных трудов. – Ульяновск, 1995 г.

[Коуэн и др., 1988] Адаптивные фильтры: Пер. с англ. / Под ред. К. Ф. Н. Коуэна и П. М. Гранта. – М.: Мир, 1988. – 392 с.

[Крашенинников, 2003] Крашенинников В.Р. Основы теории обработки изображений // Учебное пособие. – Ульяновск: УлГТУ, 2003 г. – 150 с.

[habrahabr.ru, 2010] Игровой контроллер Kinect – первое на рынке устройство для прямого трехмерного измерения – <http://habrahabr.ru/post/109337>.

УДК 681.51

БЕСПИЛОТНЫЕ ТРАНСПОРТНЫЕ СРЕДСТВА, СОВРЕМЕННОЕ СОСТОЯНИЕ И ПЕРСПЕКТИВЫ

Д.А. Добрынин (*dobr@viniti.ru*)
ВИНИТИ РАН, Москва

В докладе рассматриваются вопросы применения методов ИИ для построения беспилотных транспортных средств. Описываются современное состояние и перспективы развития интеллектуальных транспортных роботов.

Введение

Роботы прочно вошли в нашу жизнь и стали основой экономического и военного могущества развитых стран мира. В настоящее время в мире насчитывается более 1,3 млн роботов. По оценкам Международной федерации робототехники (IFR), ежегодно продается 160 тыс. роботов различного назначения.

Транспортные роботы – это беспилотные транспортные средства, обладающие элементами искусственного интеллекта (ИИ), способные перемещаться в сложной среде.

Отметим, что искусственный интеллект (ИИ) и робототехника всегда были тесно связаны друг с другом. Одним из важных направлений ИИ до сих пор считается целенаправленное поведение роботов (создание интеллектуальных роботов, способных автономно совершать операции по достижению целей, поставленных человеком) [Поспелов, 1988]. Интеллектуальный робот как машина-исполнитель должен принимать задание в общей форме, а сам робот должен обладать возможностью принимать решения или планировать свои действия в распознаваемой им неопределенной или сложной обстановке [Добрынин, 2006].

Для транспортного робота это означает, что система должна самостоятельно выбирать маршрут до намеченной цели, обходить препятствия, которые могут возникать на пути, выбирать оптимальные параметры перемещения – скорость, положение корпуса в пространстве, ускорения на виражах, а также предотвращать столкновения и аварийные ситуации.

Наибольшее развитие в настоящее время транспортные роботы получили в авиации, где они носят название БПЛА (беспилотный летательный аппарат). На основе этих технологий строятся автономные и полуавтономные робототехнические системы для разведки, наблюдения, боевого применения и транспортировки грузов.

Наземные беспилотные транспортные системы развиты менее значительно, поскольку сложность наземной среды существенно выше, чем свободный полет в воздухе. Самыми «умными» среди беспилотных систем следует считать подводные и космические роботы, поскольку особенности водной стихии и космоса накладывают серьезные ограничения на возможности телеуправления аппаратами.

За прошедшее десятилетие появилось множество разработок, ориентированных в первую очередь на создание роботов для домашнего хозяйства, где требуются достаточно сложные интеллектуальные функции. В первую очередь – это ориентация и перемещение в ограниченном пространстве с меняющейся обстановкой (предметы в доме могут менять свое местоположение), открывание и закрывание дверей при перемещении по дому.

Работы по созданию домашних роботов ведут многие крупные университеты мира, часто по заказам военных. Это связано с возможностью использования разработанных технологий в различных областях, и в первую очередь – для создания военных роботов.

Военные роботы

Наибольший прогресс в области робототехники сейчас наблюдается в сфере развития применения военных роботов. Это связано с успешным развитием программы по созданию робототизированной армии, которой занимается Defense Advanced Research Projects Agency (DARPA) – основной исследовательский центр Министерства обороны США.

По оценкам специалистов, подготовка робото-армии, состоящей из воздушных сил и пехоты, будет завершена в течение ближайших 10 лет. А к 2015 году DARPA планирует сделать одну треть транспортных средств, состоящих на вооружении США, беспилотными. Переход к полноценной робототехнической армии должен состояться к 2025 году. Мировой кризис и сокращение военного бюд-

жета США несколько затормозили этот процесс, однако об отмене речь не идет. Благодаря внедрению роботов, из военных действий исключается самый важный фактор – присутствие на поле боя живых солдат. Используя спутниковую связь, управление такой армией можно вести из любой точки мира.

Ранее вооруженные силы США уже использовали беспилотные самолеты (беспилотники, БПЛА) для оказания огневой поддержки наземным войскам в Ираке и Афганистане. В частности, беспилотники MQ-1 Predator и MQ-9 Reaper выполняли задачи по уничтожению объектов на пути наступления. Тем не менее эффективность современных беспилотных систем пока не может сравниться с пилотируемыми истребителями-бомбардировщиками A-10 или F-16. В частности, БПЛА не способны нести столько же оружия, сколько боевые самолеты.

Скорость роботизации армии США впечатляет. Например, число беспилотников в составе ВВС и Армии США увеличилось с 2000 года по 2010 г. в 136 раз: с 50 до 6800 единиц, сообщает The Air Force Times. В числе применяемых сегодня беспилотных летательных аппаратов – легкие БПЛА, запускаемые вручную, беспилотные вертолеты и тяжелые ударные машины, используемые ежедневно. Согласно принятому Пентагоном 30-летнему плану военного строительства в области авиации в период 2011–2020 финансовых годов на закупки многоцелевых беспилотных разведывательных и разведывательно-ударных систем в бюджете Министерства обороны США выделение средств будет постоянно расти и к 2020 году достигнет просто огромной по нашим меркам суммы – почти 7 млрд долларов. При этом произойдет почти семикратный рост количества имеющихся в ВС США многоцелевых беспилотных авиационных систем разведывательного и разведывательно-ударного типа.

При этом Армия США намерена к 2035 году конвертировать все имеющиеся на вооружении вертолеты в беспилотные версии с возможностью пилотирования человеком. Кроме того, военные США сейчас занимаются составлением требований к истребителю шестого поколения, который, предположительно, также станет беспилотным.

Согласно новому исследованию Teal Group, рынок беспилотников является самым быстрорастущим сектором мировой аэрокос-

мической индустрии. В ближайшие 10 лет спрос на них удвоится и вырастет с 4,4 до 8,7 млрд долл., а общий объем продаж за этот период составит 62 млрд долл. При этом 72% расходов на исследовательские, опытные и конструкторские работы придутся на долю США.

Боевые наземные робототехнические системы в настоящее время разрабатываются в виде дистанционно управляемых роботов. А вот транспортные задачи удобнее и дешевле решать с помощью беспилотных систем.

В январе 2014 года компания Lockheed Martin на полигоне Форт Худ в Техасе провела испытания системы Autonomous Mobility Applique System (AMAS), которая предназначена для превращения любого армейского грузовика в беспилотный автомобиль. Система разрабатывается в рамках 11-миллионного контракта, заключенного в 2012 году. После установки на обычный грузовик набора сенсоров и управляющих систем он сможет полностью автономно перемещаться по дорогам, в том числе и городским, объезжать препятствия, реагировать на появление пешеходов и других машин. В случае необходимости им можно будет управлять дистанционно.

Автономные грузовики смогут доставлять грузы, не рискуя жизнью водителей. Так же, как и в автомобилях, которые разрабатывает Google, беспилотный грузовик ориентируется в пространстве с помощью лидара.

Колесные средства, даже снабженные самыми изощренными системами подвески, не в состоянии сопровождать армейские подразделения в труднопроходимой горно-пустынной местности. Американская фирма Boston Dynamics уже более пяти лет занята разработкой четвероногого шагающего робота Big Dog, который должен стать верным спутником солдата в самых сложных природно-климатических условиях. Big Dog снабжен двигателем и спецсредствами, обеспечивающими ее управление, передвижение, ориентацию на местности и связь. В качестве источника энергии используется двухтактный двигатель внутреннего сгорания с водяным охлаждением мощностью в 15 л.с. Мотор приводит в движение гидронасос, каждая конечность имеет четыре гидропривода, которые приводят в движение суставы, оставляя при этом необходимые пассивные (неуправляемые) степени свободы. Бортовой

компьютер контролирует поведение боевого робота, управляет работой датчиков и обеспечивает связь с оператором. Big Dog имеет около пятидесяти сенсоров.

Робот Big Dog использует самые разнообразные режимы передвижения – может стоять, припадать к земле, передвигаться ползком (одномоментно поднимается только одна конечность), рысью (поднимаются попарно диагональные конечности), бегом, когда присутствует фаза полета, или галопом. В настоящее время Big Dog проходит тестирование в Афганистане в качестве транспортного робота сопровождения пехоты.

Израильская компания Tactical Robotics разработала беспилотник с вертикальным взлетом. Аппарат назван AirMule и сейчас проходит летные испытания. БПЛА создан для перевозки грузов, либо эвакуации раненых с поля боя. AirMule весит 770 кг, может нести полезную нагрузку весом до 640 кг, при этом может развивать скорость до 180 км/ч и летать на высоте чуть более 3600 метров. Аппарат может управляться как автопилотом, так и оператором, который находится на земле.

Активно развиваются **морские и подводные роботы**. Например, уже сегодня американская компания Hydroid выпустила порядка 200 автономных необитаемых подводных аппаратов Remus 100. Они используются многими государствами для исследований рельефа подводного дна, поиска морских мин, охраны морских военных баз.

Мировой рынок роботизированных подводных аппаратов будет проявлять высокие темпы роста в течение следующих пяти лет. Мировой рынок дистанционно управляемых роботов оценивается в \$1,2 млрд в 2014 году и, как ожидается, покажет среднегодовой темп роста в 20,11% к 2019 году. Мировой рынок полностью автономных подводных роботов оценивается в \$457 миллионов в 2014 году и, как ожидается, покажет среднегодовой темп роста в 31,95% к 2019 году.

Гражданские транспортные роботы

Беспилотные автомобили станут главным трендом предстоящего десятилетия в гражданском секторе применения транспортных роботов. В настоящее время разработки в этой области вышли в область коммерческого применения, ведущие мировые фирмы по

производству автомобилей – BMW, Nissan и др. говорят о скором появлении на рынке полностью беспилотных автомобилей.

В соответствии с прогнозом компании IHS, которая уже больше 50 лет предоставляет консалтинговые и аналитические услуги в области промышленности и технологий, к 2035 году в эксплуатации будет 54 миллиона автономных автомобилей. Годовые продажи машин-роботов составят 11,8 миллиона единиц, причем почти половина из них уже не будут иметь органов ручного управления. К 2050 году автономные автомобили полностью вытеснят обычные.

IHS предполагает появление полностью автономных автомобилей в середине 2020-х годов. Такие машины все еще будут иметь привычные органы управления – на всякий случай. Некоторые производители (например, Nissan и BMW) обещают выпустить первые подобные модели в 2020 году. Наконец, к концу 2020-х – началу 2030-х появятся машины-роботы вообще без привычных руля, педалей и рычага переключения передач. Компания Google уже выпустила такой автомобиль в тестовый режим. Лидерами по внедрению автономных автомобилей в 2035 году будут США и Канада – на них придется 29% мирового рынка, на втором месте окажется Китай с 24%, а на третьем – Европа с 20%.

Количество дорожных происшествий по мере проникновения машин-роботов снизится, уменьшится загрязнение воздуха и практически исчезнут пробки, так как машины смогут эффективнее оптимизировать движение. С другой стороны, на первый план выйдут новые риски – надежность софта и информационная безопасность. Одним из главных препятствий для внедрения автономных автомобилей станут юридические ограничения и устаревшие правила дорожного движения.

В ближайшие годы беспилотными должны стать многие грузоперевозки: и железнодорожные, и авиационные, и, главное, морские. Беспилотные морские корабли, по оценкам экспертов, просто «взорвут» отрасль контейнерных перевозок через океан, удешевив процесс на 20–30%.

Сервисные складские роботы позволяют эффективно автоматизировать работы промышленных складов. Например, Российская компания RoboCV внедряет в Калужской области на складе Samsung роботов для складских помещений X-Motion. Система

представляет из себя сервисного робота с установленной системой автопилота, которая позволяет роботу работать автономно вокруг людей в условиях склада.

Система состоит из двух частей: «серверной части», которая взаимодействует с системой управления складом, координацией распределения задач среди роботов, и навигационное решение. Роботы следуют по заранее определенным путям, но в состоянии динамично реагировать на изменение окружающей среды.

Данные постоянно передаются между транспортным средством через Wi-Fi, в то же время это позволяет работать с другими автоматизированными автомобилями в сети. Система интегрирована с системой управления складом, чтобы распределить задачи различных транспортных средств в зависимости от их положения. Это позволяет сократить время поездки и повысить общую эффективность.

Космические роботы проводят исследование космоса и других планет. Автоматические межпланетные станции, планетоходы для изучения поверхности планет солнечной системы являются по сути сложными интеллектуальными роботами, поскольку из-за больших задержек сигнала телеуправление ими в реальном времени практически невозможно.

С сентября 2010 года на космической станции работает человекоподобный робот Robonaut 2 (R2), который стал постоянным членом ее экипажа. R2 стал первой человекоподобной машиной, посланной в космос. Как сообщает пресс-релиз компании General Motors, одного из создателей «Робонавта», этот аппарат разработан для совместных с людьми действий и использования «человеческих» инструментов.

Роботы-охранники становятся все более распространенными для правительственных и корпоративных клиентов не только потому, что они снижают издержки, но и потому, что они способны улучшить параметры наблюдения и безопасности. Эти устройства автономно проводят надзор и проверку периметра на наличие злоумышленников.

С 2011 года в штате Невада роботы охраняют ядерные материалы. С марта 2014 года американские морские пехотинцы объявили, что их авиабаза в Twentynine Palms также будет охраняться мобильными роботами MDARS от General Dynamics.

MDARS являются четырехколесными дизельными беспилотниками с гидростатическим расположением руля. Грузоподъемность 500 кг. Каждый автомобиль оснащен системой видеофиксации и предупреждения в режиме реального времени. Он может работать в течение 16 часов без дозаправки на скорости до 32 км/час.

Коммерческая доставка срочных грузов с помощью летающих роботов типа мультикоптеров сулит солидные прибыли крупным компаниям, поскольку сокращает время доставки до нескольких десятков минут. В настоящее время коммерческое использование малых БПЛА ограничивается юридическими вопросами по регулированию воздушного пространства. Крупные международные компании Amazon и DHL заявили о развертывании работ в этом направлении.

Соревнования и модельные задачи

3 ноября 2007 года в местечке Викторвилль (Victorville), что в Калифорнии, успешно прошла гонка DARPA Urban Challenge – первое в мире соревнование автомобилей-роботов в городских условиях ([DARPA]). Автомобили должны были преодолеть сложный городской маршрут: и все – полностью самостоятельно, без вмешательства человека. Из-за различных проблем очень многие команды не сумели завершить гонку. Но три лидера (все, уложившиеся к тому же в 6-часовой лимит времени, определенный правилами) показали образцовую езду. Для соревнований команды оборудовали стандартные автомобили системами трехмерного зрения, лазерными дальномерами и бортовыми компьютерами. Технологии, разработанные в ходе подготовки к соревнованиям, будут использованы для разработки беспилотных автомобилей гражданского и военного назначения.

В настоящее время проходит соревнование DARPA Robotics Challenge. Цель мероприятия – создание робота, который мог бы заменить человека в зоне бедствия или катастрофы. Всего было восемь ключевых заданий: управление автомобилем, переход по кирпичам, уборка строительного мусора, открытие двери и проход в нее, подъем по лестнице, разрушение бетонной панели с помощью созданного для людей инструмента, поиск и закрытие вентилей, подключение пожарного рукава к трубе и открытие вентиля.

Конкурс проходил 20 и 21 декабря 2013 года на автостраде Хомстед-Майами, где 16 команд со всего мира постарались проявить всё своё мастерство, чтобы спроектированные ими роботы смогли выполнить серию заданий. Одним из требований к роботам было понимание простых команд вроде «Открой дверь» или «Расчисти от мусора площадку перед собой».

В финале отборочного турнира победила японская команда SCHAFT с роботом S-One, набрав 27 баллов из 32-х возможных. Примечательно, что этот небольшой японский стартап недавно приобрела корпорация Google, которая за последний год купила восемь робототехнических компаний, включая Boston Dynamics. Финал конкурса состоится в декабре 2014 года, где роботы будут бороться за главный приз DARPA – \$2 млн.

Международные соревнования EURATHLON ([EIROATHLON]) направлены на решение задач робототехники в сложных ситуациях реального мира. Испытания и сценарии выбираются с учетом важной темы исследования Европейского Союза: «восстановление безопасности в случае кризисов». Это поле исследований охватывает все основные робототехники целей, таких, как ориентация в пространстве, автономность, адаптивность и устойчивость. В идеале автономные летающие, наземные и подводные роботы должны действовать вместе, провести анализ ситуации, сбор экологических данных и выявить потенциальные опасности. При этом роботы должны передвигаться самостоятельно как в воздухе, так и на земле и в водной среде.

В 2015 году состоятся финальные соревнования EURATHLON grand challenge: воздушные, наземные и морские роботы будут действовать вместе после стихийных бедствий в ситуации, аналогичной той, что была в Фукусиме. Летающие роботы будут самостоятельно строить карту области и собирать экологические данные. Подводные аппараты будут исследовать химический разлив в море. Наземные роботы будут проводить разведку и устранять последствия аварий в зданиях. Таким образом, EURATHLON станет не только европейским, но и мировым событием в области робототехники, которое даст мощный толчок научно-техническому прогрессу в области интеллектуальных роботов.

Заключение

На сегодняшний день беспилотные транспортные системы вышли из области чисто научных разработок и стали такими же необходимыми элементами повседневной жизни, как телевидение и сотовая связь.

Дальнейшая интеллектуализация таких систем приведет не только к их широкому распространению, но как в случае с беспилотными автомобилями, к кардинальному изменению образа жизни многих людей. Именно поэтому на сегодняшний день мы наблюдаем взрывной рост разработок в области распознавания изображений, планирования траекторий, интеллектуального управления.

СПИСОК ЛИТЕРАТУРЫ

[**Добрынин, 2006**] Добрынин Д.А. Интеллектуальные роботы вчера, сегодня, завтра //X национальная конференция по искусственному интеллекту с международным участием КИИ-2006 (25–28 сентября 2006 г., Обнинск): Труды конференции. В 3-х т. Т.2. М.: Физматлит, 2006.

[**Поспелов, 1988**] Поспелов Г.С. Искусственный интеллект – основа новой информационной технологии. – М.:Наука, 1988.

[**DARPA**] DARPA official materials. <http://www.darpa.mil>.

[**IGVC**] IGVC official materials. <http://www.igvc.org>.

[**EUROATHLON**] EURATHLON robot competition <http://www.elrob.org>.

УДК 004.02

ВИЗУАЛЬНАЯ НАВИГАЦИЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ В НЕОПРЕДЕЛЕННЫХ И ИЗМЕНЯЕМЫХ УСЛОВИЯХ НАБЛЮДЕНИЯ

Н.В. Ким (*nkim2011@list.ru*)

Московский авиационный институт, Москва

Н.Е. Бодунков (*boduncov63@yandex.ru*)

Московский авиационный институт, Москва

В работе рассматриваются вопросы, связанные с организацией визуальной навигации автономным беспилотным летательным аппаратом в изменяемых и неопределенных условиях наблюдения. Предложена методика и адаптивный алгоритм распознавания площадных наземных ориентиров. Представлены результаты экспериментов и показана эффективность предложенного метода.

Введение

Одним из актуальных направлений в развитии современных беспилотных летательных аппаратов (БЛА) является разработка систем навигации, обеспечивающих возможность их автономного функционирования, в т. ч. без использования спутниковых навигационных систем (СНС).

Альтернативой СНС можно считать активно развивающиеся системы визуальной навигации. Эти системы основаны на сравнении и идентификации изображений наземных ориентиров (запомненных эталонов с известными координатами) с текущими изображениями подстилающей поверхности, получаемыми на борту БЛА.

Системы визуальной навигации являются автономными, но имеют существенные ограничения по применению, связанные со сложностью сопоставления эталонных и текущих изображений, получаемых при различных условиях наблюдения.

В [Ким и др., 2011] было предложено сопоставлять не непосредственно изображения, а их лингвистические описания. Однако реализация этого подхода требует предварительного распознавания типов подстилающих поверхностей и ориентиров.

В данной статье предлагается возможное решение этой пробле-

мы, основанное на адаптивном (к изменяемым условиям наблюдения) описании характеристик текущих изображений.

1. Адаптивный алгоритм распознавания

Процессы распознавания ориентиров производятся чаще всего с помощью корреляционно-экстремальных алгоритмов или на основе выделения и оценки признаков и использования статистических критериев (Неймана-Пирсона, Байеса и др.) с учетом возможных ошибок и потерь [Визильтер и др., 2010].

Адаптация эталонов к изменяемым условиям наблюдения в корреляционно-экстремальных системах чрезвычайно сложна, поэтому рассмотрим возможности признаковых методов.

Словарь классов распознаваемых объектов (ориентиров) обозначается так:

$$X = (x_1, x_2, \dots, x_m, \dots, x_M), \quad (1.1)$$

а вектор признаков –

$$Y = (y_1, y_2, \dots, y_n, \dots, y_N), \quad (1.2)$$

где M – количество распознаваемых классов, N – количество анализируемых признаков.

Принятие решения о принадлежности фрагмента изображения к тому или иному классу производится на основании определения области, к которой относятся признаки полученных изображений. Статистическая связь между полученными значениями признаков и классами определяется условными плотностями распределения $p(Y|x_m)$.

При решении задач распознавания наземных ориентиров возникает необходимость построения адаптивных плотностей распределения, учитывающих текущие условия наблюдения – $p(Y|x_m, Q_r)$, где $Q = (q_1, q_2, \dots, q_r, \dots, q_R)$ – вектор условий наблюдения. Под условиями наблюдения могут пониматься различные факторы, например, характеристики освещенности наблюдаемой сцены, время года, регион и т.п.

Существует несколько способов формирования таких плотностей распределения:

- на основе набора типовых условных плотностей (используется набор плотностей для каждой комбинации условий наблюдения);

- на основе использования общих физических законов (строятся модели зависимостей изменения формы плотностей распределения от изменения условий наблюдения).

В данной статье предлагается альтернативная методика построения плотностей распределения $p(Y|x_m, Q_r)$, основанная на использовании нейронечетких систем. Такой подход позволяет сократить набор типовых плотностей за счет использования экспертных правил и свойств нечетких систем, а также дает возможность их коррекции за счет использования механизмов обучения по ограниченной выборке.

Можно выделить следующие этапы методики построения предлагаемой нечеткой системы:

1. Определение общих требований (например, определение словаря классов и требований к точности распознавания);

2. Определение диапазона возможных условий наблюдения (например, наблюдение может проводиться в период «Лето» – «Осень»);

3. Определение требований к обучающей выборке по объему и качеству (объем выборки зависит от требований по точности, а качество от диапазона условий наблюдения – обучающая выборка должна входить в диапазон);

4. Получение выборки (получение эталонных изображений для каждого класса объектов и условий). Выборка может получаться путем моделирования или в результате реальной съемки;

5. Построение условных плотностей распределения $p(Y|x_m, Q_r)$ по полученным выборкам (например, на основе парзеновского окна или k-ближайших соседей);

6. Составление набора нечетких правил (помимо непосредственного формулирования правил, этот этап включает в себя формирование лингвистических переменных и их функций принадлежности).

На рисунке 1 показана структура нечеткого правила.

В нечетком правиле некоторому набору условий поставлена в соответствие конкретная плотность распределения. При этом условия записываются в виде нечетких значений, например: время года – «Лето», «Зима» и т.п., время суток – «утро», «день», «вечер», погодные условия – «дождь», «ясно» и т.д.

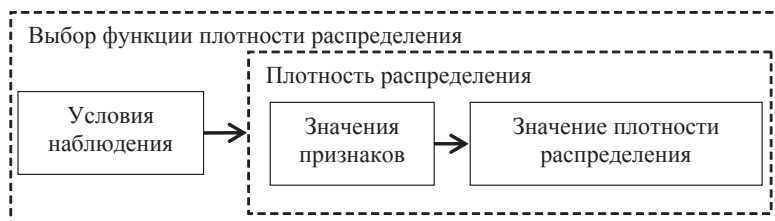


Рис. 1. Структура нечеткого правила

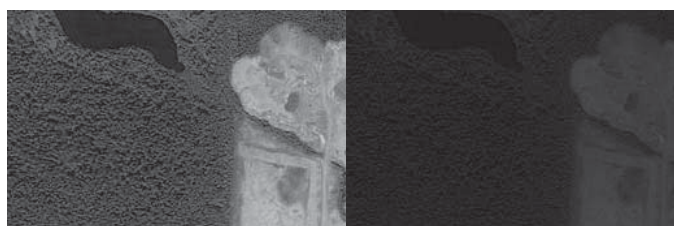
Во время работы системы в реальных условиях наблюдения может сработать одновременно несколько правил, и в итоге будет сформирована результирующая плотность распределения, усредняющая плотности, заданные в правилах.

Одной из важнейших особенностей предлагаемого подхода является возможность обучить уже сформированную систему. Во время обучения выбирается некоторое тестовое изображение. Для этого изображения производится описание условий и формируется результирующая плотность распределения. На основе этой плотности производится распознавание. Если точность распознавания неудовлетворительная, то происходит коррекция нечетких правил.

2. Пример работы системы

Рассмотрим задачу распознавания типов подстилающей поверхности: леса, поля, озера на одной сцене в разное время суток. Диапазон изменения времени примем с 12:00 до 19:00.

Изменение времени моделировалось изменением общей яркости изображения. Так было смоделировано два тестовых изображения условно в 12:00 – рисунок 2 а) и в 19:00 – рисунок 2 б).

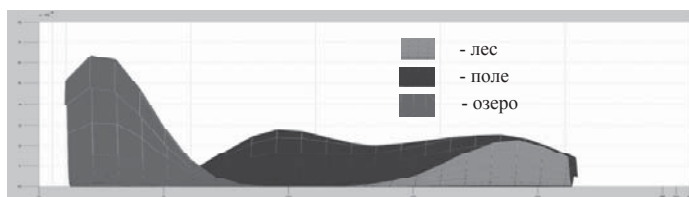


а) день

б) вечер

Рис. 2. Тестовые изображения

Для каждого условия наблюдения и для объекта (леса, озера, поля) были построены условные плотности распределения признаков математического ожидания и дисперсии яркости. На рисунках 3 а), б) и 4 а), б) представлены полученные поверхности.

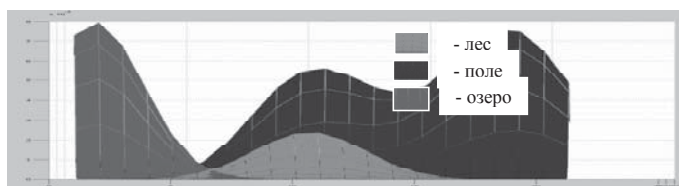


а) Математическое ожидание яркости

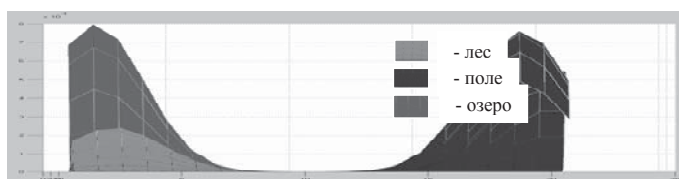


б) Дисперсия яркости

Рис. 3. Плотности распределения для изображения «День»



а) Математическое ожидание яркости



а) Математическое ожидание яркости

Рис. 4. Плотности распределения для изображения «Вечер»

Для оценки качества полученных плотностей проведем распознавание изображений обучающей выборки. Результаты распознавания приведены на рисунке 5 а) б) и в таблице 1.

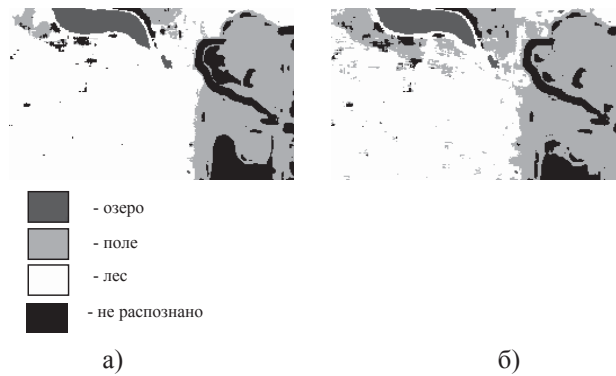


Рис. 5. Распознавание тестовых изображений

На рисунках типы поверхностей обозначены различными цветами.

Таблица 1

	Лес	Поле	Озеро
День (12:00)	90%	80%	95%
Вечер (19:00)	85%	80%	95%

В таблице 1 показан процент успешного распознавания каждого из объектов. По высоким результатам можно сделать вывод о качестве полученных плотностей распределения.

На рисунке 6 приведены примеры полученных в результате нечетких правил для двух условий: Время суток: «День» и Время суток: «Вечер».

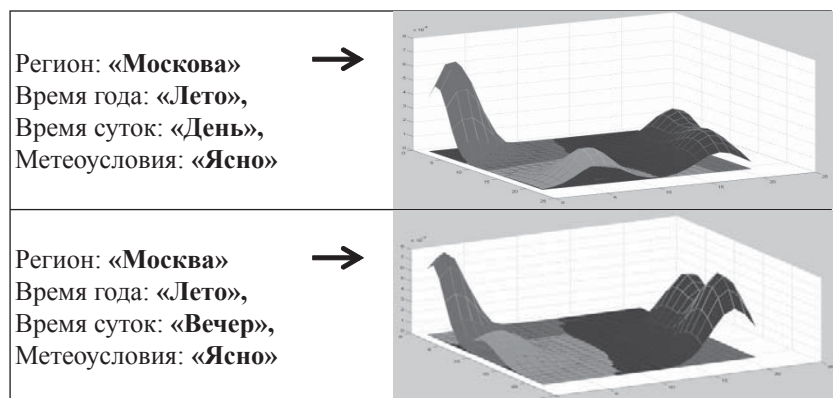


Рис. 6. Пример нечетких правил

Для оценки работоспособности полученной нечеткой системы был проведен эксперимент по распознаванию изображения на рисунке 7. Условно изображение было смоделировано для времени 16:00.

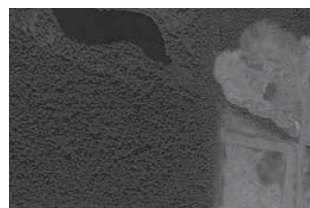


Рис. 7.
Исходное изображение

На рисунках 8 а), б) показаны результаты распознавания леса, поля и озера, при плотностях распределения вероятности обучающих изображений, а на рисунке в) при результирующей плотности распределения, полученной в результате работы нечеткой системы.

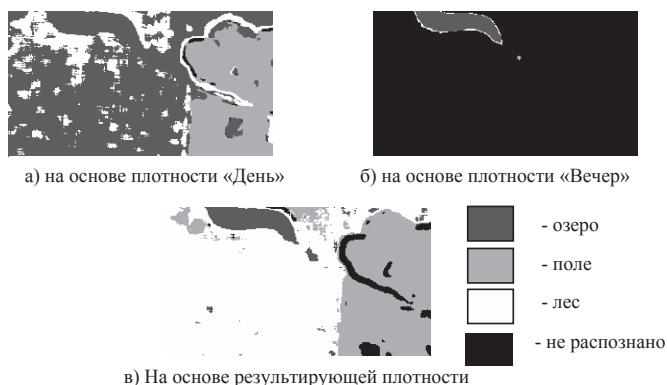


Рис. 8. Эксперименты по распознаванию с различными плотностями распределения

В таблице представлены численные значения результатов распознавания.

Таблица 2

	Лес	Поле	Озеро
Плотность «День»	10%	95%	90%
Плотность «Вечер»	5%	0%	90%
Результирующая плотность	90%	90%	90%

Таким образом, предложенная методика использования нечетких систем при распознавании площадных ориентиров показала свою эффективность в случае изменяемых условий наблюдения.

Заключение

Предложен алгоритм распознавания некоторых типов подстилающих поверхностей, адаптивный к изменению условий наблюдения.

Показано, что предлагаемый алгоритм позволяет существенно расширить условия функционирования автономных БЛА за счет повышения вероятности правильного распознавания наземных ориентиров при различных условиях наблюдения.

СПИСОК ЛИТЕРАТУРЫ

[Ким и др., 2011] Ким Н.В., Кузнецов А.Г. Автономная навигация БЛА на основе обработки и анализа видовой информации. Известия КБНЦ РАН. №1. 2011.

[Визильтер и др., 2010] Визильтер Ю.В., Желтов С.Ю. Обработка и анализ изображений в задачах машинного зрения. – М.: Физматкнига, 2010.

УДК 629.7.05

СОГЛАСОВАННОЕ УПРАВЛЕНИЕ АВТОНОМНОЙ ГРУППОЙ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Н.В. Ким (*nkim2011@list.ru*)

Московский авиационный институт, Москва

И.Г. Крылов (*Krylov_I@mail.ru*)

Московский авиационный институт, Москва

Рассматривается алгоритм формирования согласованных управлений для беспилотных летательных аппаратов при групповом поиске наземных объектов.

В настоящее время беспилотные летательные аппараты (БЛА) применяются для решения различных, как гражданских, так и военных задач [Ким и др., 2011]. В силу ограниченности возможностей одиночного БЛА все чаще рассматриваются задачи, реализуемые группой БЛА. Наиболее сложными являются задачи, в которых каждый БЛА должен выполнять свою подзадачу согласованно с другими БЛА.

В работе [Ким и др., 2011] был предложен сценарий поиска (обнаружения и оценки координат) наземных малоразмерных подвижных объектов группой БЛА, функционирующей в автономном режиме (без участия человека-оператора), и общий алгоритм формирования управлений для каждого БЛА. Было показано, что ключевым является автоматическое распределение заданной общей целевой задачи (ЦЗ) на отдельные подзадачи для каждого БЛА. В рамках рассматриваемых задач наблюдения это означает разделение области поиска на подобласти, обследуемые отдельными БЛА с помощью размещенных на них систем наблюдения (СН). При этом необходимо обеспечить согласованность их действий и оптимизацию процесса с учетом заданного критерия, например, продолжительности поиска

$$T_{\Pi} = \max T_m, \quad (1)$$

где T_m – время обследования m -м БЛА соответствующей подобласти, $m \in M$.

Вероятность правильного обнаружения $P_{по}$ обычно задается в виде

$$P_{по} \geq P_{тр}, \quad (2)$$

где $P_{тр}$ – заданная вероятность правильного обнаружения.

Распределение области поиска с учетом (1), (2) и технических характеристик БЛА является сложной научно-технической задачей. В частности, возможность обнаружения объектов поиска с заданной достоверностью (2) определяется множеством факторов, одним из которых является загромождение этих объектов другими объектами, например, зданиями, деревьями и пр.

Таким образом, необходимым (но недостаточным) условием для обнаружения или распознавания объектов поиска является их наблюдаемость, зависящая от положения СН, размеров объектов поиска и структуры рельефа обследуемой местности. Наблюдаемость объектов с БЛА выше на открытых участках местности и уменьшается при обследовании городских районов или леса. Чем выше наблюдаемость объектов, тем на большем расстоянии могут прокладываться траектории полетов БЛА, соответственно, повышая производительность поиска.

Карты области наблюдения, на которых обозначены расстояния между траекториями полета БЛА, с учетом типов подстилающей поверхности и характеристик объектов поиска будем называть *картами наблюдаемости*.

В представленной работе предлагается подход, обеспечивающий предварительное распределение заданной ЦЗ поиска для группы БЛА вертолетного типа (рис. 1) с учетом наблюдаемости объектов на различных подстилающих поверхностях.

В соответствии с представленным алгоритмом на 1-м этапе ЦЗ, передаваемая ведущему группы БЛА, представляется в формализованном виде, позволяющем в бортовых вычислителях БЛА планировать и реализовать дальнейшие действия группы, подгрупп и отдельных БЛА. В ЦЗ входят: вид требуемого поиска, описание объектов поиска, описание области поиска, требования к оценке эффективности поиска, дополнительные требования, а также атрибуты объекта поиска, атрибуты области поиска (границы различных поверхностей и их тип), условия поиска и уникальные требования к прохождению маршрута. Предполагается, что имеется карта области, а границы подстилающих поверхностей и их типы определены.

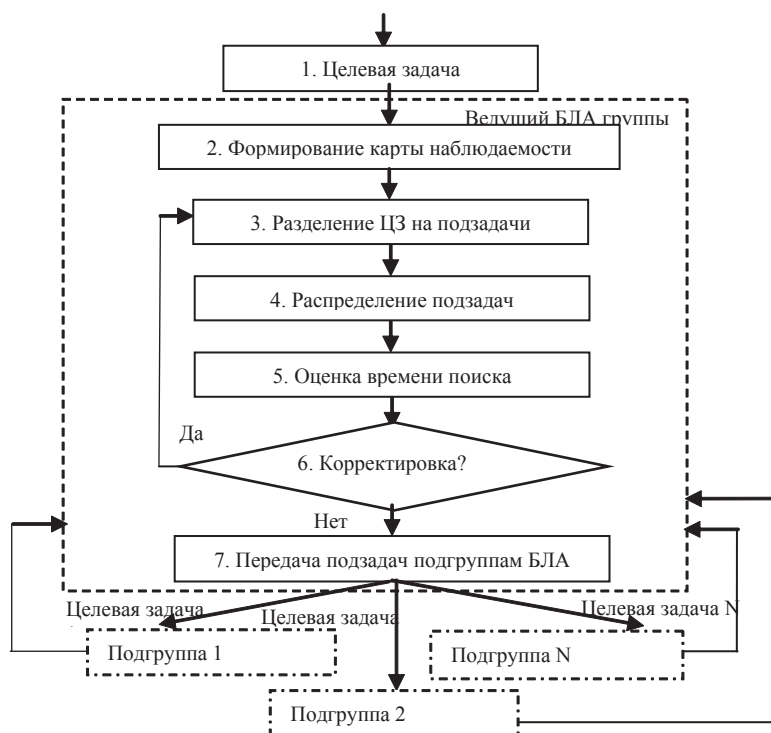


Рис. 1. Структура общего алгоритма управления

На следующем этапе формируется карта наблюдаемости области поиска. При этом будем считать, что для конкретных условий наблюдения и типов подстилающих поверхностей можно определить некоторые унифицированные расстояния между траекториями полета БЛА. На основе моделирования процессов поиска были получены оценки расстояний между траекториями полета БЛА, обеспечивающие достаточно подробное обследование (в соответствии с требованием (2)) различных подстилающих поверхностей при поиске малоразмерных объектов (размером не более 3x3 м).

Так, при высоте полета 200 м БЛА вертолетного типа для гарантированной наблюдаемости ($P_{гр} = 0,95$) малоразмерного объекта расстояния между траекториями должны быть не более:

- 10 м – для леса;
- 20 м – для населенного пункта;

100 м – для поля, водной поверхности.

На основе этих данных может быть сформирована карта наблюдаемости, в которую входят показания к построению маршрута над разными видами поверхностей, например, лес, город, поле и т.д.

На следующем этапе (рис. 1, блок 3) область поиска разделяется на подобласти, которые распределяются между отдельными БЛА в группе (рис. 1, блок 4).

Для расчета времени T_m строятся маршруты облета подобластей с учетом технических характеристик и типов БЛА

$$T_m = L_m / V_{\text{блa}}$$

где L_m – длина траектории полета БЛА в m -й подобласти, $V_{\text{блa}}$ – скорость полета m -го БЛА.

В качестве примера рассмотрим процедуру распределения подзадач между 4-мя БЛА вертолетного типа при поиске малоразмерных объектов в области, представленной на рисунке 2.



Рис. 2. Исследуемая область

Вариант 1. Будем считать, что в области поиска присутствуют 3 типа подстилающих поверхностей: лес, населенный пункт, поле (водная поверхность). Таким образом, в данном примере для каждой поверхности при планировании маршрутов БЛА необходимо выбирать соответствующее расстояние между траекториями.

С учетом типов поверхностей данной области поиска и результатов моделирования была получена соответствующая карта наблюдаемости (рисунок 3).

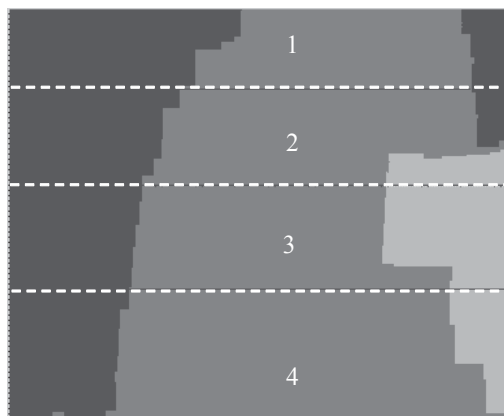


Рис. 3. Карта наблюдаемости

На рисунке 3 присутствуют области, окрашенные 3-мя цветами: лес – черным (Ч), поле – серым (С), населенный пункт – светло-серым (СС).

Общая площадь области С составляет 577269 м², площадь области СС составляет 102468 м², площадь области Ч составляет 303576 м².

С учетом приведенных выше требуемых расстояний между траекториями получим:

- длина маршрута над областью С равна 5772 м,
- длина маршрута над областью СС равна 5123 м,
- длина маршрута над областью Ч равна 30358 м.

Общая длина маршрута составляет $L_0 = 41254$ м.

Пусть скорости БЛА равны, тогда наилучшим будет вариант равномерного распределения длины маршрута между БЛА

$$L_m = \frac{L_0}{4} = 10313 \text{ м.}$$

Границы распределенных областей показаны на рисунке 3 штриховыми линиями.

При скорости $V_{\text{блa}} = 15 \text{ м/с}$ получим время обследования (вариант 1) всей области поиска

$$T_{\text{п1}} = T_m = 688 \text{ с.}$$

Вариант 2. При разделении области поиска между БЛА на равные площади без учета наблюдаемости объектов на различных подстилающих поверхностях получим в области:

- 1 длина маршрута – 12860 м, время наблюдений – 857 с;
- 2 – 10606 м, – 707 с;
- 3 – 9624 м, – 642 с;
- 4 – 7853 м, – 524с.

Согласно (1), время обследования области поиска для варианта 2 равно

$$T_{П2} = T_m = 857 с.$$

Экономия времени в варианте 1 по сравнению с вариантом 2 составляет 25%.

Заключение

Предложен алгоритм формирования согласованных управлений БЛА при групповом поиске наземных объектов. Показано, что учет наблюдаемости объектов на различных типах подстилающих поверхностей позволяет существенно повысить производительность наблюдений.

СПИСОК ЛИТЕРАТУРЫ

[Ким и др., 2011] Ким Н.В., Кузнецов А.Г. Поиск объектов на основе анализа наблюдаемой ситуации // Сб. тезисов докладов на научном семинаре «Системы технического зрения» ИКИ РАН / М., 2011.

[Ким и др., 2011] Ким Н.В., Крылов И.Г., Лебедев А.В. Согласованное применение группы БЛА // Материалы шестой Всероссийской научно-практической конференции «Перспективные системы и задачи управления» и третьей Молодежной школы-семинара «Управление и обработка информации в технических системах». – Таганрог: Изд-во ТТИ ЮФУ, 2011, с. 11–14.

УДК 004.896

ИНТЕЛЛЕКТУАЛЬНАЯ ОБРАБОТКА ИНФОРМАЦИИ В СИСТЕМЕ ТЕХНИЧЕСКОГО ЗРЕНИЯ

М.О. Корлякова (*mkorlyakova@yandex.ru*)

П.К. Новиков (*bobersuperzy_b@mail.ru*)

Е.Ю. Прокопов (*evgen-prokopov@rambler.ru*)

А.Ю. Пилипенко (*gamerjd@mail.ru*)

Московский государственный технический университет, Калужский филиал, Калуга

В работе представлены подходы к применению интеллектуальных технологий обработки информации при создании систем технического зрения для бортовых навигационных комплексов. Рассмотрено решение основных задач обработки изображений и стереореконструкции с использованием нейросетевой аппроксимации и классификации. Приведены примеры реализации отдельных этапов обработки информации в нейросетевом базисе для бортовых систем технического зрения.

Введение

Автономные мобильные системы все шире включаются в решение практических задач перемещения грузов и патрулирования территорий в естественных средах, которые не подготовлены специально для работы роботов. Широко известны проекты беспилотных автомобилей корпорации Google, конкурс автоботов и роботов, проводимый DARPA [Stavens et al.], система ASIMO (Honda) и другие проекты в области робототехники. Общим для них является направленность на работу на обычных городских или природных площадках, где необходимо определять свое положение и его изменение. Это заставляет использовать для бортовых навигационных систем объекты и детали местности. В качестве средств формирования модели сцены принято использовать весь комплекс доступных датчиков: эхолокацию, лазерные дальнометры, системы технического зрения (СТЗ) и т.п. Однако в целом ряде случаев применение активных систем наблюдения должно быть ограничено. Например, при решении задач работы в агрессивных средах, где любое внешнее воздействие может спровоцировать ухудшение си-

туации или проинформировать противника о проникновении на его территорию. Тогда единственным способом формирования модели среды становится система технического зрения, которая пассивно «рассматривает» сцену.

Работа СТЗ является вычислительно сложной задачей, требующей значительных ресурсов, которые должны быть предоставлены вычислительной системой мобильного объекта. Кроме того, точность определения положения, передвигающегося в естественной среде транспортного средства, зависит от размера кадра получаемого с камер СТЗ. Незначительное увеличение размера кадра приведет к существенному росту времени обработки, а небольшое уменьшение размера кадра быстро снижает точность. Необходимо обеспечить баланс точности и скорости обработки, поскольку бортовые вычислительные системы подвижных объектов имеют жесткое ограничение ресурсов, но требуют максимально возможной точности. Таким образом, для решения задачи ориентации и навигации с использованием СТЗ на борту автономной мобильной платформы необходимо искать пути снижения вычислительной сложности задач обработки и анализа изображений.

Применение нейросетевых технологий обработки информации на борту движущегося транспортного средства или в системах навигации роботов позволит повысить производительность бортовой СТЗ, не снижая ее точности. Построение нейросетевой системы технического зрения для решения задачи навигации и ориентации в естественной среде рассмотрим на примере анализа скорости и перемещения мобильной платформы.

Бортовая система технического зрения

Основной подход к решению задачи анализа скорости и перемещения состоит в исследовании снимков сцены, получаемых с нескольких камер СТЗ в процессе движения мобильной платформы. В качестве платформы будем использовать автономное наземное транспортное средство (автомобиль, поезд, мобильный робот и т.п.). Скорость движения платформы не превышает 30 км/ч. В самом простом случае используем стереопару камер и рассматриваем левый и правый кадры, чтобы по проекциям объектов построить их истинное положение. Камеры СТЗ связаны между собой и отка-

либрованы. Скорость съемки, как правило, значительно превышает темп обработки информации в СТЗ. Поэтому будем пропускать часть снятых кадров, пока идет обработка последней пары для поиска перемещения и скорости. Общая схема обработки имеет следующий вид:

- получение пары кадров в момент времени t_i и t_{i+1} на борту мобильного объекта;
- выделение характерных объектов сцены и их особых точек (углы, линии, эллипсы) на изображениях в моменты времени t_i и t_{i+1} , т.е. нахождение одинаковых объектов, ракурс которых изменился от одного кадра к другому в бортовой системе зрения;
- определение параметров размещения СТЗ относительно объектов сцены в моменты времени t_i и t_{i+1} , т.е. решение задачи определения трехмерных координат объекта по его проекционной модели – стереорекострукция;
- вычисление перемещения платформы на основании информации о координатах СТЗ в пространстве;
- определение скорости относительно моментов времени t_i и t_{i+1} .

Для решения задачи поиска перемещения в идеальном случае достаточно знания пространственных координат 4-х точек сцены в исходном положении и в текущем. Однако точное определение координат объектов невозможно, что значительно усложняет поставленную задачу [Кокарева и др., 2012]. Этап калибровки СТЗ реализуется до старта мобильной платформы, а, следовательно, не влияет на скорость решения задачи на маршруте. Другой вопрос, что калибровка неидеальна, а ошибка калибровки включается в общую ошибку определения координат. Помимо этого, ошибка определения истинной координаты складывается из влияния размытия точек (функция размытия точки зависит от аппаратной части СТЗ), ошибки алгоритма поиска особых точек, которая неизбежно возникает на этапе выделения объектов, ошибки стереорекострукции и сопоставления координат объектов сцены на этапе вычисления перемещения. Процедуры обработки изображений и стереорекострукции существенно отличаются с точки зрения формальной математической постановки задачи. Однако они могут быть реализованы за счет обучения нейронных сетей, что позволит строить алгоритм общего решения в однородной форме.

Преобразование и предобработка изображений в нейросетевых системах являются наиболее отработанной частью общей схемы обработки, приведенной выше. Существует множество примеров использования ИНС в системах технического зрения для разных прикладных задач как на уровне предобработки изображений [Sartin et al., 2014], так и на уровне классификации объектов [Ozbay et al., 2005; Lari et al., 2011].

В рамках поставленной задачи, на этапе предобработки, с точки зрения жесткого ограничения ресурсов и времени для бортовой навигационной системы, наиболее сложным является этап поиска особых точек. Рассмотрим подходы к его нейросетевой реализации в форме сети Хопфилда, которая обеспечит распознавание в условиях зашумленных образцов, и сети радиальных нейронов (RBF сети), которая основана на поиске ближайшего шаблона.

Поиск особых точек изображения

В качестве входных данных использовали базу примеров углов и прямых краев, описанных в форме бинарной матрицы. Исходное изображение подвергается контурному анализу с применением фильтра Собеля [Гонсалес и др., 2005], а поиск особых точек осуществляется только для областей, прилегающих к контурам. Пример шаблонов для поиска особых точек приведен на рис. 1.

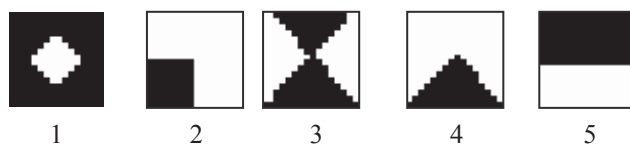
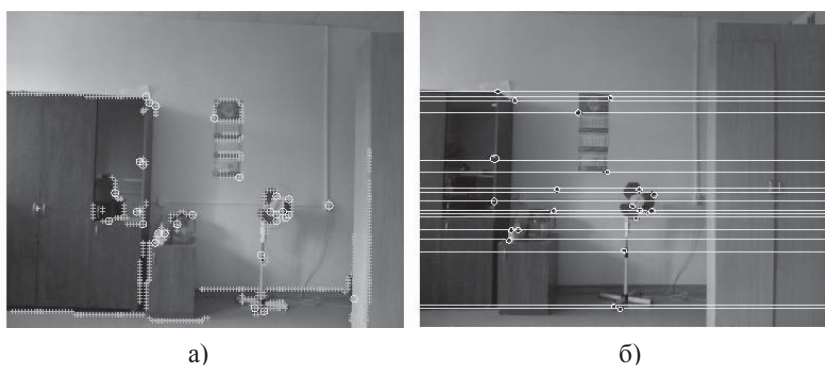


Рис. 1. Примеры шаблонов для поиска углов (1-4) и краев (5)

В процессе распознавания точку кандидата с ее окрестностью размером 16x16 пикселей подают на вход нейросети и оценивают результат на выходе. Эксперименты показали, что обучение сети Хопфилда позволяет определить незначительное количество угловых точек и приводит к значительному (в 3–4 раза по сравнению с RBF сетью) увеличению времени обработки кадров. Сеть радиальных нейронов сформирована за счет принудительного размещения нейронов в точках, соответствующих шаблонам. Она определяет

ближайший шаблон и возвращает вектор кода типа образца. Кроме того, число угловых областей, выделяемых таким способом, значительно больше, чем у сетей Хопфилда. Таким образом, в результате экспериментальной проверки в качестве метода выделения особых точек принята модель сравнения с шаблонами в сети радиальных нейронов. Пример выбора особых точек (белый «круг») из точек кандидатов («крест») для левого кадра стереопары приведен на рис. 2. а.



а) б)
Рис. 2. Поиск особых точек на левом кадре а)
и перенос их на правый кадр б)

Каждая особая точка имеет описание растровой области $I_{16 \times 16}$ пикселей, описание градиента изображения $R_{16 \times 16}$ и тип шаблона, который для нее был определен. Это позволяет значительно сократить этап сопоставления особых точек левого и правого кадров, т.к. сопряженные особые точки должны иметь близкий тип, похожие описания областей $R_{16 \times 16}$ и $I_{16 \times 16}$ и лежать в одной эпиполярной плоскости [Кокарева и др., 2012]. Следует отметить, что главным достоинством разработанного алгоритма является его быстроедействие, по сравнению с методом перебора и анализа дескрипторов, но нет инвариантности к поворотам и масштабам.

Стереореконструкция для особых точек изображения

Следующий сложный этап – решение задачи стереореконструкции, где из-за ошибок определения положения сопряженных особых точек левого и правого кадров не существует точного решения. В [Кокарева и др., 2012] предложен механизм нейросетевой аппрок-

симации трехмерных координат точки по ее двум стереопроjectionам. Обучение нейросети провели в виртуальной среде с использованием модели СТЗ, полученной на этапе калибровки. Для аппроксимации координат использовали сеть типа многослойный перцептрон с нелинейным скрытым слоем и 3-мя линейными выходными нейронами, которую обучали на 10000 примеров. Входы нейронной сети: координаты проекций объекта, а выходы – его трехмерные координаты. Все примеры были сформированы в виртуальной модели сцены. Эффект достигается за счет переноса основной вычислительной сложности системы на этап обучения нейронной сети, который занимает довольно значительное время (3–5 часов с учетом построения выборки примеров). Кроме того, формальная математическая модель стереорекострукции позволяет работать в любом диапазоне дальностей и ошибки координат будут пропорциональны дальности. Нейросетевой подход требует ограничения диапазона глубины сцены при обучении и решении задачи и показывает резкий рост ошибок за пределами этого диапазона. Результаты экспериментов по вычислению координат для виртуальной модели и реальной сцены показывают, что ошибка стереорекострукции составляет 1–2% дальности объекта. Скорость получения ответа от нейросети $3 \cdot 10^{-4}$ с. и значительно превосходит методы поиска оптимальных координат (0,02 сек.). Результаты стереорекострукции сцены с рис. 2 приведены на рис. 3.а.

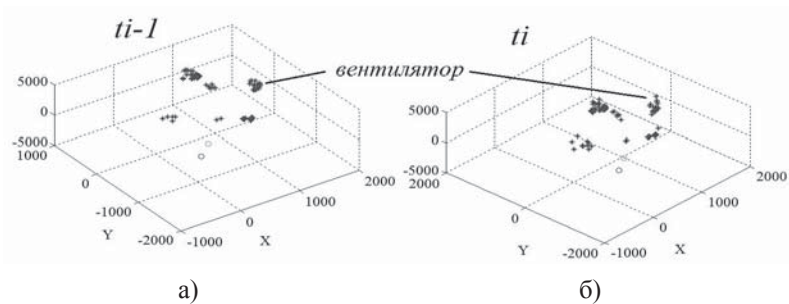


Рис. 3. Стереорекострукция особых точек сцены для момента t_i а) и особых точек сцены для момента t_{i+1} б)

На этапе подготовки стереорекострукции особые точки сцены отсеиваются на основе эвристического алгоритма для получения наиболее достоверного положения объектов сцены.

Поиск перемещения бортовой системы технического зрения

Стереореконструкцию особых точек проводим с кадрами стереопары для момента t_i и t_{i+1} (рис. 3 а и б). Результаты стереореконструкции для момента t_i и t_{i+1} образуют множество координат V (положение точек наблюдаемой сцены в момент времени t_i) и V_{new} (положение тех же точек сцены в момент времени t_{i+1}). Перемещение объектов сцены относительно камеры или камеры относительно объектов сцены можно описать уравнением вида:

$$\begin{aligned} M \cdot V &= V_{new} \\ M &= V_{new} \cdot V^{-1} \end{aligned}$$

где M – матрица переноса-поворота мобильного объекта с СТЗ на борту.

Идеального решения, в силу существования ошибок при определении V и V_{new} , не существует. Построим нейросетевую аппроксимацию этого решения, а поскольку нейронные сети устойчивы к шуму, то возможно получить качественное решение задачи поиска параметров M при ошибках определения координат 1–2% от дальности. В результате экспериментов было показано, что наиболее качественное решение получается за счет разделения общего перемещения на поворот и перенос в следующем виде: $M = R \cdot T$, R – поворот, T – перенос. Причем поиск параметров поворота осуществляем не по V и V_{new} , а по трем ортам направлений наибольшей дисперсии точек сцены для их центрированных представлений. Тогда частные ошибки отдельных точек из множеств V и V_{new} значительно меньше влияют на общий результат.

Для поиска параметров R использовали многослойный перцептрон с одним нелинейным скрытым слоем. Орты множеств V и V_{new} строим за счет использования метода главных компонент [Ивашина и др., 2012]. Входы нейронной сети содержат координаты ортов центрированного множества точек V_{new} , размещенные в системе координат, заданной ортами множества V . Выходы сети – параметры матрицы поворота R . Все примеры формируем в виртуальной сцене в виде идеального значения R , сопоставленного идеальному значению ортов. К сожалению, устойчивое состояние ортов сцены можно гарантировать только при значительных (несколько тысяч точек) объемах V и V_{new} . На рис. 4 приведены ошибки поворота от

числа доступных точек для ошибки стереорекострукции 20 мм, на средней глубине 2 м.

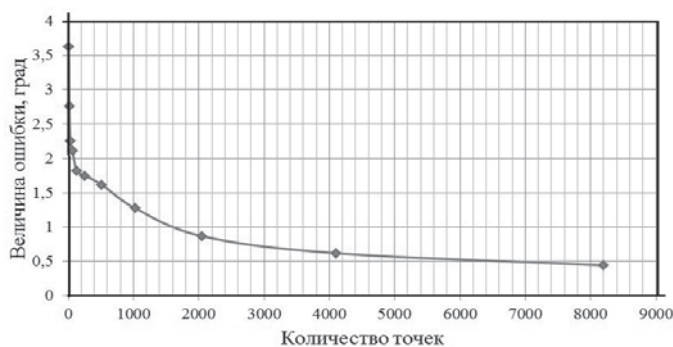


Рис. 4. Ошибка определения поворота от числа доступных точек объектов сцены

Ошибка поворота становится менее 1 гр. при практически недоступных объемах данных, но 100–200 точек сцены позволяют получать не более чем 1.5–2 гр. ошибки при общем повороте до 90 гр.

Перенос T при известных поворотах определяется за счет анализа V_m, V_{m_new} средних для множеств V и V_{new} , как $T = V_m \cdot R - V_{m_new}$. Причем, поскольку T зависит от средних по V и V_{new} , то влияние их индивидуальных ошибок практически исключается, а общая ошибка зависит в основном от ошибки определения поворота и составляет 30–40 мм при глубине сцены 2 м. Время обработки первой стереопары, при размере каждого изображения 1280x1024 пикселей, составляет 42 с., а каждой последующей – 30 с. при моделировании в среде MatLab. Для решения аналогичной задачи без использования нейронных сетей пришлось затратить 120–200 с. при эквивалентном уровне ошибок.

Заключение

Результаты моделирования нейросетевой СТЗ показали принципиальную эффективность подхода к решению задач навигации и ориентации в бортовых системах автономных мобильных объектов. Получены вычислительно простые решения задач стереорекострукции и поиска параметров перемещения. Ошибки традици-

онной и интеллектуальной модели СТЗ сопоставимы по уровням, а время обработки существенно меньше. Следует отметить, что области эффективной работы нейросетевой модели ограничены, что потребует определения типа среды, где перемещается мобильная платформа.

Однородность решений позволит построить более простую аппаратную и программную реализацию модели в бортовой вычислительной системе робота или автономного транспортного средства.

СПИСОК ЛИТЕРАТУРЫ

[Гонсалес и др., 2005] Вудс Р., Гонсалес Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.

[Ивашина и др., 2012] Ивашина Е.А., Корлякова М.О., Пилипенко А.Ю. Нейросетевое восстановление координат сцены в стереоскопической системе технического зрения// Гибридные и синергетические интеллектуальные системы: теория и практика: материалы 1-го международного симпозиума/ под ред. проф. А.В.Колесникова. – Калининград: Изд-во БФУ им.И.Канта, 2012. – С. 163–173.

[Кокарева и др., 2012] Кокарева Е.А., Корлякова М.О., Пилипенко А.Ю. Решение задачи стереорекострукции в нейросетевом базисе // XIV Всероссийская научно-техническая конференция «Нейроинформатика 2012»: Сборник научных трудов. В 3-х частях. Ч.1. – М.: НИЯУ МИФИ, 2012. – С. 160–169.

[Lari et al., 2011] Lari Z. and Ebadi H., Automated Building Extraction from High-Resolution Satellite Imagery using Spectral and Structural Information Based on Artificial Neural Networks// International conference on object extraction, vol. 21, no.21. pp. 23–27.

[Ozbay et al., 2005] S. Ozbay and E. Ercelebi, Automatic Vehicle Identification by Plate Recognition A Neural Network// World Academy of Science, Engineering and Technology, vol. 9, (2005), pp. 222–225.

[Sartin et al., 2014] Sartin Maicon A. and da Silva Alexandre C. R., Evaluation of Image Segmentation and Filtering With Ann in the Papaya Leaf// International Journal of Computer Science & Information Technology (IJCSIT), Vol. 6, February 2014, pp. 47–58.

[Stavens et al.] Stavens D. and Thrun S. A Self-Supervised Terrain Roughness Estimator for Off-Road Autonomous Driving // <http://arxiv.org/1206.6872>.

УДК 629.7.058

ЗАДАЧИ УПРАВЛЕНИЯ И ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ПРИ ПРИМЕНЕНИИ ГРУПП БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

А.А. Липатов (*my-post-box@hotmail.ru*)

Д.А. Миляков (*from_fn@mail.ru*)

ОАО «Концерн «Вега», Москва

В работе проведен анализ ряда задач управления и информационного обеспечения, возникающих в связи с применением групп беспилотных летательных аппаратов (БЛА). Рассмотрены способы решения задач оптимального централизованного управления траекторией и пространственным построением группы БЛА, а также выявления таких групп с использованием неопределенной информации количественного и качественного характера.

Введение

За последнее десятилетие интерес к беспилотным летательным аппаратам (БЛА) заметно возрос. По мнению большинства экспертов, БЛА являются наиболее перспективным направлением развития авиации. Это объясняется не только тем, что массовое производство и применение БЛА оказываются дешевле и проще, чем пилотируемых летательных аппаратов (ЛА), но и тем, что некоторые типы БЛА способны решать задачи, недоступные пилотируемым ЛА, к примеру, задачи ближней разведки в условиях плотной городской застройки, в помещениях и т.п. При этом повреждение БЛА в ходе выполнения задачи приводит лишь к финансовым потерям и не угрожает жизни пилота-оператора.

Долгое время БЛА находили только военное применение. При этом принятие в ряде стран концепции бесконтактных сетевых войн [Верба, 2009] привело к усилению роли БЛА в качестве средств как нападения, так и разведки, и защиты. В этом смысле БЛА должны восприниматься как летающие роботы с относительно высокими показателями автономности, т.е. способные решать ряд задач без непосредственного участия оператора [Каляев и др., 2009].

Однако в последние годы растет потребность и в БЛА гражданского назначения, например, для мониторинга состояния нефте- и

газопроводов, лесных массивов, водоемов, сбора метеорологических данных и т.д. Также БЛА находят применение в поисковых и спасательных операциях. Таким образом, развитие БЛА гражданского назначения имеет большие перспективы.

Наиболее важными задачами, решаемыми БЛА как военного, так и гражданского назначения, являются:

- многоспектральный мониторинг местности, под которым понимается ее обследование с использованием оптических, инфракрасных, радиолокационных и других датчиков информации;
- картографирование местности;
- формирование мобильных коммуникационных сетей и др.

Решение указанных задач требует перехода от применения одиночных БЛА к использованию групп однотипных и разнотипных БЛА. Это связано с тем, что мониторинг объектов, имеющих большую площадь или протяженность, для одиночных БЛА является весьма затруднительным. Кроме того, ограничения по массе целевой нагрузки БЛА не позволяют реализовать многоспектральный мониторинг местности с помощью одиночного аппарата. Применение БЛА в экстремальных условиях и, тем более, в условиях организованного противодействия нередко приводит к их повреждению или полному выходу из строя. В случае группового применения БЛА это может затруднить решение поставленной задачи, но не приведет к прекращению ее выполнения.

Кроме того, в условиях конфликта увеличение количества БЛА в группе может приводить к созданию эффекта «роя», при котором их количество превышает пропускную способность средств противодействия [Канащенко и др., 2002], что обеспечивает группе БЛА существенное тактическое преимущество. В связи с этим важной задачей становится автоматическое выявление и сопровождение групп взаимодействующих воздушных объектов (ВО), в частности БЛА.

Необходимо подчеркнуть, что переход к групповому применению БЛА приводит к возникновению ряда специфических задач управленческого и информационного характера.

Среди наиболее сложных задач управления следует отметить:

- задачи траекторного управления группами БЛА, осложненные большой размерностью моделей состояния и нелинейным ростом размерности алгоритмов управления;

- задачи формирования и поддержания требуемого пространственного построения группы БЛА.

Для информационного обеспечения групповых действий БЛА в условиях конфликта большое значение имеет решение задачи заблаговременного выявления и последующего сопровождения враждебных групп БЛА при возможности проявления эффекта «роя».

Целью настоящего доклада является выявление актуальных задач управления группами БЛА и информационного обеспечения этого процесса, а также рассмотрение и разработка способов решения некоторых из них.

1. Задачи управления группами БЛА

Особенностью задач траекторного управления группой БЛА является то, что помимо управления отдельными БЛА необходимо обеспечить согласованность их действий с действиями других БЛА группы.

В настоящее время известны два основных вида стратегий управления группами БЛА [Каляев и др., 2009]: стратегии централизованного и децентрализованного управления, которые имеют ряд разновидностей.

Следует, однако, подчеркнуть, что независимо от способа синтеза управления группой БЛА на основе классической теории оптимального управления необходимо использовать уравнения высокой размерности, а в качестве критериев оптимальности – минимум квадратичных функционалов качества с высокоразмерными матрицами штрафов за точность управления и расход энергии управляющих сигналов.

Эти особенности приводят к тому, что в процессе формирования сигналов управления группой БЛА необходимо совместно решать системы уравнений очень высокой размерности. Это явление, известное как «проклятие размерности» [Ройтенберг, 1992], затрудняет использование классических алгоритмов оптимизации (в постановке Летова-Калмана, Красовского, Понтрягина) при решении задач управления большими группами БЛА.

Одним из примеров снижения размерности данной задачи может служить алгоритм оптимального централизованного управления группой БЛА [Меркулов и др., 2012]. Снижение размерности достигается следующими способами:

- декомпозицией вектора состояния группы БЛА на подвекторы, соответствующие отдельным объектам или их составным частям;

- заменой классического функционала качества с использованием общей квадратичной формы, отвечающей за точность функционирования, на модифицированный функционал с суммой квадратичных форм для каждого подвектора;

- последовательным индивидуальным управлением каждым БЛА с учетом состояния других БЛА в группе вместо одновременного управления всеми БЛА.

Другой класс задач управления группами БЛА связан с необходимостью формирования и поддержания требуемого пространственного построения группы. Эти задачи также называют строевыми задачами [Каляев и др., 2009].

В общем виде данная задача формулируется следующим образом: при заданных исходных и требуемых положениях БЛА группы в пространстве, допустимой погрешности построения, ограничениях на положения БЛА и на их возможные перемещения требуется переместить БЛА из исходного положения в требуемое.

В зависимости от того, каким образом задается требуемое положение БЛА, общую задачу можно разбить на два класса.

В задачах первого класса требуемые положения БЛА группы задаются в виде абсолютных координат в пространстве.

Задачи второго класса требуют поддерживать определенное взаимное положение БЛА в строю.

В свою очередь, задачи первого класса разбиваются на два подкласса. К первому подклассу относятся задачи, в которых координаты требуемого положения определены для каждого конкретного БЛА группы. Задачи данного подкласса характерны для групп, состоящих из небольшого числа разнородных БЛА, которые должны занимать строго определенные позиции в строю [Каляев и др., 2009]. Примерами задач такого типа служат транспортировка крупных тел (БЛА выстраиваются по периметру перемещаемого тела), сбор проб (БЛА занимают строго заданные места для забора образцов), охрана протяженных объектов (БЛА группы перемещаются на контрольные точки) и т.п. Решение задачи сводится к перемещению каждого БЛА из начальной точки в заданную точку с обходом/облетом препятствий и предотвращением столкновений.

Задачи второго подкласса связаны с группами из однотипных БЛА. В этом случае не имеет большого значения, какой из БЛА группы займет то или иное положение в строю, а описание требуемого положения группы содержит координаты требуемых точек, но не содержит информации о том, какому БЛА какая точка соответствует [Каляев и др., 2009].

На практике задача такого типа встречается в тех же приложениях, что и предыдущая. Решение такой задачи разбивается на два этапа. На первом этапе решается задача о назначениях, в результате чего каждому БЛА ставится в соответствие определенное требуемое положение, а затем задача сводится к предыдущей.

Существенно большую сложность имеют задачи второго класса, в которых первоочередное значение имеет обеспечение заданного взаимного положения БЛА в строю. В частности, положение БЛА относительно друг друга может задаваться в виде матрицы относительных расстояний между ними. Следует отметить, что такие задачи чаще встречаются на практике, чем задачи первого класса.

Особенность данного класса задач заключается в том, что строевому расположению, задаваемому в виде относительных расстояний между БЛА, может отвечать бесконечное множество различных наборов точек в пространстве. Поэтому в первую очередь необходимо определить, каким образом найти такое конкретное множество точек пространства, отвечающее заданному строевому расположению, переход в которое всех БЛА группы позволяет минимизировать пройденные ими расстояния и, как следствие, временные и энергетические затраты.

В свою очередь, данная задача также может быть разбита на два подкласса. К первому подклассу относятся те задачи, в которых задано как требуемое строевое расположение в виде относительных расстояний между БЛА, так и соответствие между каждой позицией и БЛА, который должен ее занять. Такой тип задач характерен для групп разнотипных БЛА.

Второй подкласс составляют задачи, в которых требуемое строевое расположение задано в виде относительных расстояний между БЛА, а соответствие между конкретными БЛА и позициями в строю не задано. Подобные задачи актуальны для групп однотипных БЛА.

2. Задача выявления групп воздушных объектов с учетом неопределенности их координат состояния

Задача выявления групп (группирования) взаимодействующих ВО актуальна в связи с широким распространением групповых действий пилотируемых и беспилотных летательных аппаратов.

Автоматическое группирование ВО в первую очередь призвано облегчить анализ тактической обстановки операторам, управляющим группами БЛА. Кроме того, результаты группирования ВО могут использоваться для решения задачи целераспределения в группе БЛА [Горощенко, 2000].

Известные способы решения этой задачи основаны на оценке расстояния между ВО в пространстве их координат состояния (КС), которые могут быть как количественными (координаты в пространстве, курс, скорость), так и качественными (класс, тип, государственная принадлежность) [Горощенко, 2000; Кирсанов и др. 1996]. Однако эти способы не учитывают возможной неопределенности значений количественных и качественных КС ВО, связанной с ошибками их оценивания и неполнотой доступной информации об объектах.

В связи с этим были разработаны способы группирования ВО с учетом их неопределённых количественных КС [Верба и др., 2014]. В данном докладе предлагается способ, который позволяет проводить группирование ВО с использованием как количественных, так и качественных КС.

Под количественными КС понимаются координаты в пространстве, курс и скорость объектов. Из числа качественных КС (признаков) рассматриваются государственная принадлежность и класс ВО.

При решении задачи группирования ВО p представляется набором своих КС $\langle x_1^p, x_2^p, \dots, x_k^p, x_{k+1}^p, \dots, x_n^p \rangle$, где $\langle x_1^p, x_2^p, \dots, x_k^p \rangle$ – его количественные координаты, а $\langle x_{k+1}^p, x_{k+2}^p, \dots, x_n^p \rangle$ – качественные признаки.

Каждая количественная КС x_i^p , $i = 1, k$ может быть представлена интервалом с нижней границей $x_{ин}^p = \hat{x}_i^p - \Delta_i^p$ и верхней границей $x_{ив}^p = \hat{x}_i^p + \Delta_i^p$, где $\hat{x}_i^p = a_i^p + \xi_i^p$ – текущая оценка i -й КС ВО p (число), a_i^p – ее фактическое значение, ξ_i^p – ошибка ее измерения, $\Delta_i^p = 3\sigma_i^p$, σ_i^p – среднеквадратичное отклонение случайной величины ξ_i^p .

Для качественных признаков x_i^p , $i = \overline{k+1, n}$ известны конечные множества их возможных значений A_i , а текущие оценки их значений представлены множествами $B_i^p \subseteq A_i$. При $|B_i^p| = 1$ значение x_i^p известно точно, а при $|B_i^p| = |A_i|$ оно полностью неопределено.

Предполагается, что расстояние между парой непосредственно взаимодействующих ВО p и q в пространстве их количественных КС не превышает некоторого априорно заданного порога [Кирсанов и др., 1996], а качественные признаки этих объектов, такие, как государственная принадлежность и класс, совпадают.

Проверка пары ВО p и q на наличие непосредственного взаимодействия между ними осуществляется путем попарного сравнения значений их одноименных КС. Для каждой пары КС (x_i^p, x_i^q) , $i = \overline{1, n}$ задано условие $Cr_i(x_i^p, x_i^q)$ непосредственного взаимодействия пары ВО p и q в одной группе. При этом с учетом неопределенности значений x_i^p, x_i^q , $i = \overline{1, n}$ возможны три результата проверки условия $Cr_i(x_i^p, x_i^q)$ [Верба и др., 2014]:

- условие *не выполняется*;
- условие *возможно выполняется*;
- условие *достоверно выполняется*.

Заключение о непосредственном взаимодействии пары объектов p и q в составе одной группы делается по следующим правилам.

Если существует $i \in \{1, 2, \dots, n\}$, для которого $Cr_i(x_i^p, x_i^q)$ не выполняется, то объекты p и q *непосредственно не взаимодействуют* в составе группы.

В противном случае, если существует $i \in \{1, 2, \dots, n\}$, для которого $Cr_i(x_i^p, x_i^q)$ возможно выполняется, то объекты p и q *возможно непосредственно взаимодействуют* в составе группы.

В противном случае ВО p и q *достоверно непосредственно взаимодействуют* в составе группы.

Условие $Cr_i(x_i^p, x_i^q)$ для неопределенных количественных КС x_i^p, x_i^q , $i = \overline{1, k}$ требует, чтобы различие их значений не превышало априорно заданного порога ε_i [Кирсанов и др., 1996] и для интервальных КС сформулировано следующим образом [Верба и др., 2014].

Если $|\hat{x}_i^p - \hat{x}_i^q| > \varepsilon_i + \Delta_i^p + \Delta_i^q$, то $Cr_i(x_i^p, x_i^q)$ *не выполняется*.

Если $|\hat{x}_i^p - \hat{x}_i^q| \leq \varepsilon_i + \Delta_i^p + \Delta_i^q$ и $|\hat{x}_i^p - \hat{x}_i^q| > \varepsilon_i - \Delta_i^p - \Delta_i^q$, то $Cr_i(x_i^p, x_i^q)$ *возможно выполняется*.

Если $|\hat{x}_i^p - \hat{x}_i^q| \leq \varepsilon_i - \Delta_i^p - \Delta_i^q$, то $Cr_i(x_i^p, x_i^q)$ *достоверно выполняется*.

Одним из возможных условий принадлежности пары ВО к одной группе является совпадение их качественных признаков (например, государственной принадлежности или класса). Для неопределенных качественных признаков $x_i^p, x_i^q, i = k+1, n$ условие $Cr_i(x_i^p, x_i^q)$ формализуется следующим образом.

Если $B_i^p \cap B_i^q = \emptyset$, то $Cr_i(x_i^p, x_i^q)$ не выполняется.

Если $|B_i^p| > 1$ и (или) $|B_i^q| > 1$ и $B_i^p \cap B_i^q \neq \emptyset$, то $Cr_i(x_i^p, x_i^q)$ возможно выполняется.

Если $|B_i^p| = |B_i^q| = 1$ и $B_i^p = B_i^q$, то $Cr_i(x_i^p, x_i^q)$ достоверно выполняется.

С использованием приведенных правил группирования ВО может быть построен граф их взаимодействия, вершины которого соответствуют ВО, а ребра могут быть двух типов и отображают возможное или достоверное непосредственное взаимодействие между парой ВО. Путем выявления связанных компонент этого графа могут быть выделены возможные и достоверные группы ВО [Верба и др., 2014].

Таким образом, предложенный способ группирования ВО дает следующие преимущества:

- облегчение работы операторов, управляющих группами БЛА, за счет автоматизации процесса группирования;
- оценка степени достоверности состава выявленных групп ВО за счет разделения их на возможные и достоверные;
- увеличение времени сопровождения групп ВО за счет того, что возможные группы выявляются раньше, чем достоверные, а при кратковременном расхождении ВО достоверные группы не снимаются сразу с сопровождения, а переводятся в класс возможных.

Заключение

Проведенный анализ состояния и тенденций развития тематики управления группами БЛА позволяет сделать следующие выводы.

Актуальными являются задачи оптимального управления траекторией полета группы БЛА и ее пространственным построением для различных стратегий группового управления, а также задачи информационного обеспечения групповых действий БЛА, включая выявление, оценку состава и сопровождение групп ВО. К настоя-

шему времени предложены решения отдельных разновидностей указанных задач, достоинствами которых являются, в частности, снижение вычислительной сложности алгоритмов оптимального группового управления и способность работы с неопределенной количественной и качественной информацией.

СПИСОК ЛИТЕРАТУРЫ

[Верба, 2009] Верба В.С., Поливанов С.С. Организация информационного обмена в сетевых боевых операциях // Радиотехника. 2009. №8.

[Каляев и др., 2009] Каляев И.А., Гайдук А.Р., Капустян С.Г. Методы и модели коллективного управления в группах роботов. – М.: ФИЗМАТЛИТ, 2009.

[Канащенков и др., 2002] Канащенков А.И., Меркулов В.И., Самарин О.Ф. Облик перспективных бортовых радиолокационных систем. Возможности и ограничения. – М.: ИПРЖР, 2002.

[Ройтенберг, 1992] Ройтенберг Я.Н. Автоматическое управление. – М.: Наука, 1992.

[Меркулов и др., 2012] Меркулов В.И., Харьков В.П., Шамаров Н.Н. Оптимизация коллективного управления группой беспилотных летательных аппаратов // Информационно-измерительные и управляющие системы. – 2012, №7.

[Горощенко, 2000] Горощенко Л.Б. Методы координированного наведения и атаки несколькими истребителями группы самолетов противника // Полет. – 2000, №6.

[Кирсанов и др., 1996] Кирсанов А.П., Сорвенков О.С., Сузанский Д.Н. Методы обработки радиолокационной информации при сопровождении компактных групп воздушных объектов // Радиотехника. – 1996, №10.

[Верба и др., 2014] Верба В.С., Липатов А.А., Федисов А.Н. Выявление групп воздушных объектов с учетом неопределенности их координат состояния // Успехи современной радиоэлектроники. – 2014, №1.

УДК 62-503.5

**ИССЛЕДОВАНИЕ ДВИЖЕНИЯ ПРЫГАЮЩЕГО РОБОТА,
ОСНАЩЕННОГО КРЫЛЬЯМИ
С ИЗМЕНЯЕМОЙ ГЕОМЕТРИЕЙ¹**

О.Г. Локтионова (*teormeh@inbox.ru*)

Л.Ю. Ворочаева (*mila180888@yandex.ru*)

А.В. Ворочаев (*sasha-vorochaev@yandex.ru*)

Юго-Западный государственный университет, Курск

В работе рассматривается прыгающе-летающий робот, крылья которого обеспечивают планирование устройства во время этапа полета, разработана математическая модель движения аппарата, особое внимание уделено взаимодействию крыльев с окружающей средой.

Введение

В современном мире разработке конструкций и исследованию движения прыгающих роботов уделяется большое внимание российских и зарубежных ученых, что объясняется хорошей проходимостью таких устройств при движении по неровной и пересеченной местности [Armour и др., 2007], [Hyon и др., 2002], [Stoeter и др., 2002], [Weiss, 2001]. Это в перспективе позволит внедрять их в различные области промышленности и хозяйства, например, использовать для мониторинга местности, проведения поисковых и спасательных работ, а также разведывательных мероприятий в военных целях. Использование прыгающими роботами крыльев, посредством которых обеспечивается их планирование во время прыжка, расширяет функциональные возможности аппаратов за счет увеличения дальности прыжка и высоты, с которой может осуществляться прыжок, последнее достигается путем снижения скорости устройства при падении на поверхность [Burdick и др., 2003], [Roberts и др., 2008]. Все это говорит об актуальности темы исследования.

¹ Работа выполнена при финансовой поддержке РФФИ (проект №14-08-00581).

1. Описание прыгающего робота, оснащенного системой крыльев

В работе рассматривается робот, состоящий из прыжкового и летательного модулей, трехмерная модель которого приведена на рис. 1. Прыжковый модуль образован звеньями 1 – 4, причем звено 1 представляет собой стопу, звенья 2 и 3 образуют ногу, а звено 4 является корпусом. На звеньях 4 и 2 установлены приводы вращательного движения 6 и 7 соответственно, посредством которых обеспечивается поворот звеньев 3 и 1 относительно звеньев 4 и 2. На звене 2 закреплен привод поступательного движения 8, обеспечивающий перемещение звеньев 2 и 3 относительно друг друга. Корпус представляет собой каркас в форме параллелепипеда. Летательный модуль состоит из крыльев 5, поворот относительно корпуса и раскрытие которых осуществляется приводом 9. Движение устройства по поверхности происходит за счет прыжков и раскрытия крыльев в наивысшей точке прыжка. Последнее позволяет увеличить дальность полета робота и уменьшить скорость его падения, что, в свою очередь, защищает устройство от повреждения при прыжке с большой высоты.

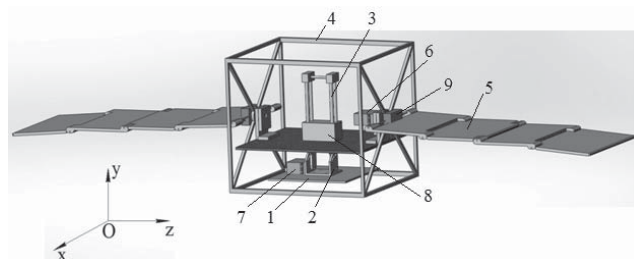


Рис. 1. Трехмерная модель прыгающе-летающего робота

2. Математическая модель движения робота во время полета

Для разработки математической модели прыгающе-летающего аппарата рассмотрим схему, показанную на рис. 2.

Будем считать, что все звенья устройства являются абсолютно твердыми телами, причем звенья 1 – 3, 5 представляют собой стержни длинами l_i , $i=1-3, 5$, а звено 4 – прямоугольник $A_1A_2A_3A_4$

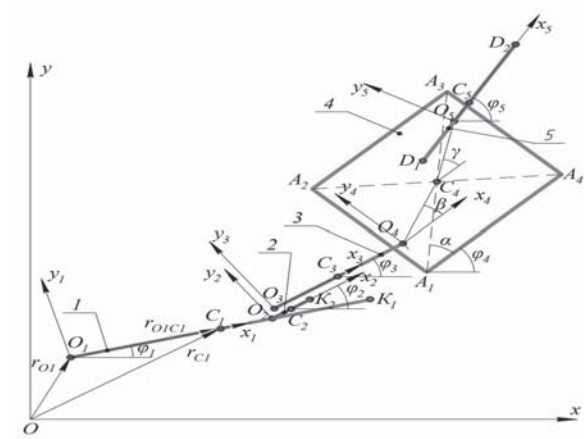


Рис. 2. Расчетная схема прыгающе-летающего робота

размерами $2ax2b$. Центры масс $m_i, i=1-5$, звеньев сосредоточены в центрах их симметрии – точках C_i . Положим, что масса крыльев очень мала, поэтому учитывать ее в математической модели не будем. Длина ноги задается расстоянием l_{23} между точками O_2 и O_4 , положение последней в корпусе определяется расстоянием l_4 и углом β . Положение точки O_5 закрепления крыла в корпусе задается расстоянием C_4O_5 и углом γ .

Методика прыжка и математические модели этапов были описаны в работах [Волкова и др., 2013], [Яцун и др., 2013a], [Яцун и др., 2013b]. Отличием данной модели является наличие крыльев, поэтому особое внимание уделено влиянию среды на прыжок. Поэтому в статье записаны дифференциальные уравнения, описывающие поведение робота только на этапе полета.

Во время этого этапа аппарат движется с отрывом от поверхности, центр масс корпуса перемещается вдоль осей Ox и Oy , а также корпус может вращаться относительно его центра масс. Звенья 3 и 5 жестко зафиксированы в корпусе устройства, звенья 2 и 1 – между собой. Звенья ноги 1 и 2 во время полета втягиваются в корпус при помощи силы F_{32} , формируемой приводом 7, до тех пор, пока длина ноги не станет наименьшей ($l_{23}=l_{23}^{\min}$), это соответствует случаю, когда время втягивания ноги в корпус меньше времени полета (рис. 3, а). Крылья 5 используются для планирования аппарата после достижения им высшей точки прыжка.

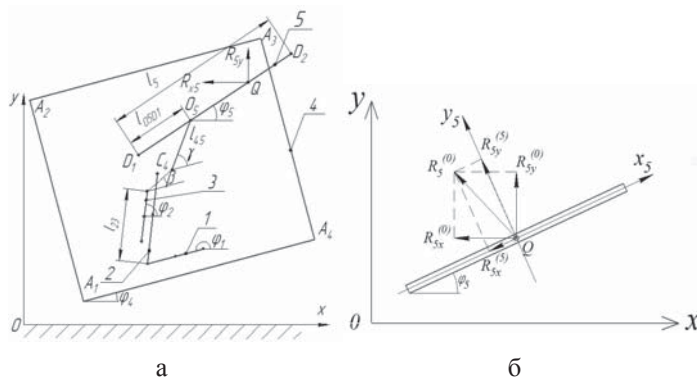


Рис. 3. а – Расчетная схема прыгающе-летающего робота на этапе полета, б – Схема приложения аэродинамической силы к крылу

На этом этапе движение робота описывается четырьмя дифференциальными уравнениями по обобщенным координатам x_{C4} , y_{C4} , φ_4 , l_{23} :

$$a_{11}\ddot{x}_{C4} + a_{13}\ddot{\varphi}_4 + a_{14}\ddot{l}_{23} + b_{13}\dot{\varphi}_4^2 + c_{11}\dot{l}_{23}\dot{\varphi}_4 = R_{5x}, \quad (2.1)$$

$$a_{22}\ddot{y}_4 + a_{23}\ddot{\varphi}_4 + a_{24}\ddot{l}_{23} + b_{23}\dot{\varphi}_4^2 + c_{22}\dot{l}_{23}\dot{\varphi}_4 = -(m_1 + m_2 + m_3 + m_4)g + R_{5y}, \quad (2.2)$$

$$a_{33}\ddot{\varphi}_4 + a_{31}\ddot{x}_4 + a_{32}\ddot{y}_4 + a_{34}\ddot{l}_{23} + b_{33}\dot{\varphi}_4^2 + c_{33}\dot{l}_{23}\dot{\varphi}_4 = \\ = (m_1 + m_2 + m_3)g \frac{l_4}{2} \cos(\varphi_4 + \beta) + M_5, \quad (2.3)$$

$$a_{44}\ddot{l}_{23} + a_{41}\ddot{x}_4 + a_{42}\ddot{y}_4 + a_{43}\ddot{\varphi}_4 + b_{43}\dot{\varphi}_4^2 = F_{32} - (m_1 + m_2)g \sin\varphi_2, \quad (2.4)$$

где a_{ij} , b_{ij} , c_{ij} , $i=1-4, j=1-4$ – коэффициенты.

В качестве точки приложения аэродинамической силы будем рассматривать точку Q крыла, положение которой определяется следующим образом:

$$l_{D1Q} = \frac{(3Al_5^2 + 4Bl_5 + 6C)l_5}{4Al_5^2 + 6Bl_5 + 12C}, \quad (2.5)$$

где A , B и C – коэффициенты.

Проекция аэродинамической силы, приложенной к точке Q , вычисляются по формулам (рис. 3, б):

$$R_{5x} = \frac{C_R \rho S_5^{O5x5}}{2} \cdot \dot{y}_Q^{(5)} \left| \dot{y}_Q^{(5)} \right| \cos\varphi_5, \quad (2.6)$$

$$R_{5y} = \frac{C_R \rho S_5^{O_5 x_5}}{2} \cdot \dot{y}_Q^{(5)} \left| \dot{y}_Q^{(5)} \right| \sin \phi_5, \quad (2.7)$$

где ρ – массовая плотность воздуха, C_R – безразмерный коэффициент аэродинамической силы, $S_5^{O_5 x_5}$ – площадь крыла вдоль оси, $O_5 x_5$, $\dot{y}_Q^{(5)}$ – проекция скорости точки Q на ось $O_5 y_5$.

Аэродинамические силы определяются из условий:

$$R_{5x} = \begin{cases} R_{5x}, & \text{если } \dot{y}_{C4} < 0 \\ 0, & \text{если } \dot{y}_{C4} \geq 0 \end{cases}, R_{5y} = \begin{cases} R_{5y}, & \text{если } \dot{y}_{C4} < 0 \\ 0, & \text{если } \dot{y}_{C4} \geq 0 \end{cases}. \quad (2.8)$$

Момент, создаваемый аэродинамическими силами, имеет вид:

$$M_5 = R_{5x} (l_{45} \cos(\gamma + \phi_4) - (l_{O_5 D_1} - l_{D_1 Q}) \cos \phi_5) - R_{5y} (l_{45} \sin(\gamma + \phi_4) - (l_{O_5 D_1} - l_{D_1 Q}) \sin \phi_5). \quad (2.9)$$

3. Исследование влияния крыльев на траекторию полета

Разработанная математическая модель позволила провести исследование влияния аэродинамических сил и момента на траекторию движения центра масс корпуса прыгающе-летающего робота. Моделирование проводилось для робота массой 0,1 кг с размерами корпуса $a=0,1$ м, $b=0,1$ м и длиной крыла $l_{D_1 D_2}=0,2$ м. Массовая плотность воздуха $\rho=1,2$ кг·с²/м⁴, коэффициент аэродинамической силы $C_R=0,6$. Рассмотрим случай, когда крыло закреплено в корпусе в точке C_5 , совпадающей с центром масс корпуса ($l_{45}=0$ м), вращения робота в полете при этом не происходит, положение точки Q на крыле остается неизменным и равным $l_{D_1 Q} = 0,5l_5$ (рис. 4).

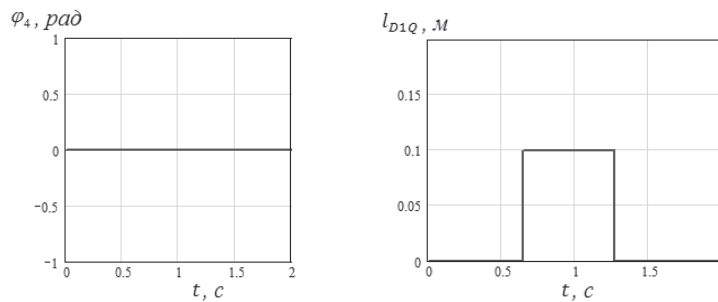
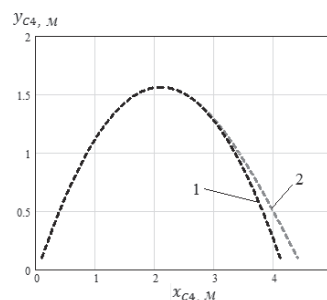


Рис. 4. Графики зависимостей: а – $\phi_4(t)$, б – $l_{D1Q}(t)$

На рис. 5 представлены графики траекторий движения робота при наличии крыльев (кривая 2) и без них (кривая 1), по которым видно, что дальность прыжка робота с крыльями возрастает примерно на 8%.

Рис. 5. Траектория движения устройства:
1 – без крыльев;
2 – при наличии крыльев



На рис. 6 показано влияние площади крыльев S_5 на траекторию движения устройства.

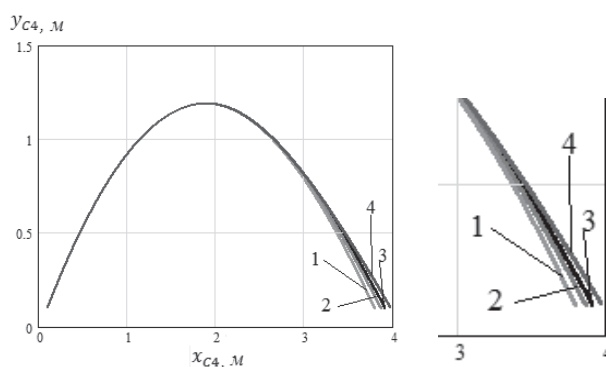
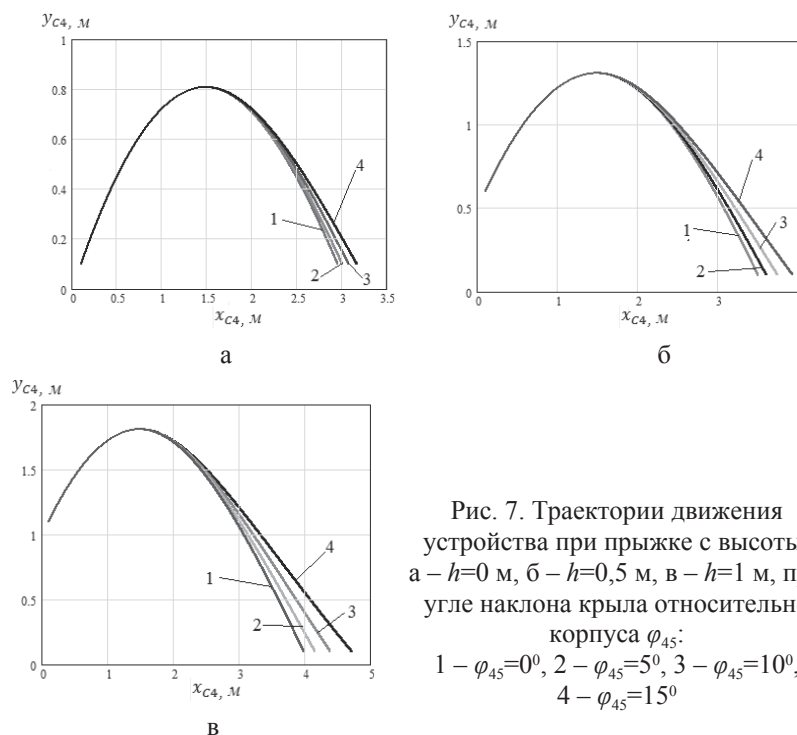


Рис. 6. Траектории движения устройства при: 1 – $S_5=0,1 \text{ м}^2$, 2 – $S_5=0,16 \text{ м}^2$, 3 – $S_5=0,2 \text{ м}^2$, 4 – $S_5=0,25 \text{ м}^2$

Из полученного графика видно, что чем больше площадь крыльев, тем больше длина прыжка. При прыжке с нулевой высоты влияние площади крыльев на дальность прыжка сравнительно незначительное, это обусловлено тем, что аппарат не успевает набрать достаточную вертикальную скорость, чтобы значение подъемной аэродинамической силы позволило существенно увеличить длину прыжка. Так, увеличение площади крыла в 2,5 раза позволяет добиться возрастания дальности полета на 0,25 м, что составляет 6%

от первоначальной длины прыжка. Из этого можно сделать вывод, что при прыжке с нулевой высоты использование крыльев нецелесообразно.

Для исследования влияния угла φ_{45} наклона крыла относительно корпуса на траекторию движения устройства было проведено моделирование прыжка для различных углов φ_{45} с трех разных начальных высот (рис. 7).



На основании анализа полученных траекторий установлено, что с увеличением угла наклона крыла относительно корпуса длина прыжка возрастает. Причем эта зависимость наиболее хорошо просматривается при прыжках не с нулевой высоты. Так, при прыжке с нулевой высоты приращение длины прыжка при $\varphi_{45}=15^\circ$ по сравнению с $\varphi_{45}=0^\circ$ равно 0,3 м, что составляет 10% от исходной дальности прыжка, с высоты $h=0,5$ м длина прыжка увеличивается на 0,5 м (15%), с высоты $h=1$ м – на 0,8 м (20%).

Заключение

В работе рассмотрен прыгающе-летающий робот, разработана его математическая модель. Особое внимание уделено аэродинамическому взаимодействию робота с окружающей средой во время этапа полета. Проведено исследование влияния крыльев на траекторию движения устройства и дальность полета, а именно их площади и угла наклона к корпусу при прыжке с различных высот. Установлено, что наиболее эффективно крылья могут использоваться для прыжков робота с высоты, существенно превышающей его габаритные размеры при ненулевом относительном угле между крыльями и корпусом.

СПИСОК ЛИТЕРАТУРЫ

[Волкова и др., 2013] Волкова Л.Ю., Яцун С.Ф. Изучение влияния положения точки закрепления ноги прыгающего робота в корпусе на характер движения устройства // *Нелинейная динамика*. – 2013. – Т. 9. – №2. – С. 327–342.

[Яцун и др., 2013а] Яцун С.Ф., Волкова Л.Ю. Моделирование движения многозвенного прыгающего робота и исследование его характеристик // *Известия РАН. Теория и системы управления*. 2013. – № 4. – С. 137–149.

[Яцун и др., 2013б] Яцун С.Ф., Волкова Л.Ю., Ворочаев А.В. Исследование режимов разгона четырехзвенного прыгающего аппарата // *Известия Волгоградского государственного технического университета*. 2013. – № 24 (127). – С. 86–92.

[Armour et al., 2007] Armour R., Paskins K., Bowyer A., Vincent J. F. V., McGill W. Jumping robots: a biomimetic solution to locomotion across rough terrain // *Bioinspiration and Biomimetics Journal*. 2007. No. 2. P. 65–82.

[Burdick et al., 2003] Burdick J., Fiorini P. Minimalist jumping robot for celestial exploration // *The International Journal of Robotics Research*, 2003, Vol. 22(7), p. 653–674.

[Hyon et al., 2002] Hyon S., Mita T. Development of a biologically inspired hopping robot – Kenken // *Proc. of the IEEE International Conference on Robotics and Automation*. Stanford. 2002. P. 3984–3991.

[Roberts et al., 2008] Roberts J. F., Zufferey J.-C., Floreano D. Energy management for indoor hovering robots // *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008, p. 1242–1247.

[Stoeter et al., 2002] Stoeter S. A., Rybski P. E., Papanikolopoulos N. Autonomous stair-hopping with scout robots // *IEEE/RSJ international conference on intelligent robots and systems*. 2002. Vol. 1. p. 721–726.

[Weiss, 2001] Weiss P. Hop. . . hop. . . hopbots!: designers of small, mobile robots take cues from grasshoppers and frogs // *Science News*. 2001. №. 159. 88 p.

УДК 531.1

МОДЕЛЬ, БАЗОВЫЕ ТРАЕКТОРИИ, НЕЙРОКОНТРОЛЛЕР ДЛЯ МУЛЬТИРОТОРНОГО РОБОТА

В.Е. Павловский (*vlpavl@mail.ru*)
ИПМ им. М.В. Келдыша РАН, Москва
А.В. Савицкий (*as@ttorr.ru*)
МГУ им. М.В. Ломоносова, Москва

В данной работе построена теоретико-механическая модель мультироторного робота – квадрокоптера с учетом основных механических эффектов. Были изучены базовые траектории движения, в том числе взлет-посадка и полет по прямой. Проведено численное моделирование. На основании проведенных расчетов строится нейроконтроллер для квадрокоптера.

Введение

В течение последних лет тема беспилотных летательных аппаратов (БПЛА) получает все большее распространение. Особый интерес вызывают мультикоптерные роботы – устройства, приводимые в движение роторами (как правило, их количество – от 1 до 8). Случай одного ротора – это вертолет, четырех – квадрокоптер – объект внимания данной работы.

Целью настоящей работы являются изучение динамики мультироторного летательного аппарата, построение базовых алгоритмов управления и численное моделирование. На основании моделирования строится нейроконтроллер для квадрокоптера. Одно из его назначений – преодоление проблемы дефицита управлений, т.к. квадрокоптер это система с их недостатком, система имеет (в полете) 6 степеней свободы и только 4 управления, которыми являются скорости вращения винтов.

1. Описание модели

В качестве квадрокоптера рассмотрим плоское твердое тело, состоящее из корпуса и четырех пропеллеров (рис. 1). Корпусом будем считать два перпендикулярных стержня длины $2l$, пересекающихся в центрах A . Центр масс квадрокоптера массы m находится

в точке A . Четыре пропеллера закреплены в точках C_1, C_2, C_3, C_4 на расстоянии l от точки A .

Пусть $\{Oxyz\}$ – правая инерциальная система координат, причем Oz соответствует вертикальному направлению. Положение центра масс тела определяется вектором $\zeta=(x, y, z)$. Пусть $\{Ae_1e_2e_3\}$ – правая система координат, жестко связанная с квадрокоптером. Выберем ось Oe_1 по направлению одной из «рук» летательного аппарата, ось Oe_2 – по направлению другой «руки».

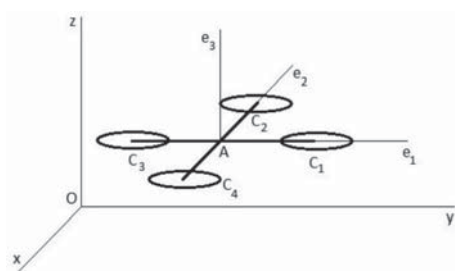


Рис. 1. Концептуальная модель квадрокоптера

1.1. Основные предположения

1) Из-за высокой угловой скорости лопастей несущих винтов будем рассматривать их как диски, а не как отдельные вращающиеся лопасти. Эти диски называются дисками несущего винта.

2) Будем рассматривать только умеренные маневры. Следовательно, мы пренебрегаем упругостью лопастей несущего винта, а диск несущего винта считается бесконечно жестким и бесконечно тонким. Диск все время вращается в плоскости, перпендикулярной оси Oe_3 .

3) Подъемная сила каждого диска создает силу поступательного движения u_i , приложенную в точке C_i и все время направленную вдоль оси Oe_3 .

4) Управление ориентацией задается с помощью трех независимых вращающихся моментов $\{\Gamma_1, \Gamma_2, \Gamma_3\}$ вокруг каждой из трех осей системы координат $\{Ae_1e_2e_3\}$. Эти вращающиеся моменты приложены непосредственно к корпусу и не порождают никаких сил поступательного движения, связанных с вторичными аэродинамическими эффектами или упругостью лопастей несущих винтов.

5) Полагая, что линейная скорость квадрокоптера и его собственные вращения существенно меньше скорости вращения вин-

тов, будем считать, что воздух оказывает сопротивление только на лопасти несущих винтов, создавая тем самым вращательные моменты трения.

б) Векторы угловых скоростей винтов, расположенных в точках C_1 и C_3 , сонаправлены вектору e_3 , другие два – противоположно направлены.

1.2. Вычисление Лагранжиана

Ориентация корпуса квадрокоптера задается с помощью трех углов Крылова $v = (\varphi, \psi, \theta)$.

Соответствующую матрицу в неподвижной системе координат обозначим за R .

Кроме того, чтобы учесть динамику несущих винтов, введем углы поворота лопастей $\dot{\gamma}_1, \dot{\gamma}_2, \dot{\gamma}_3$ и $\dot{\gamma}_4$. От них будут зависеть управляющие моменты.

Таким образом, обобщенные координаты для квадрокоптера имеют вид: $q = (x, y, z, \varphi, \psi, \theta)$.

Обозначим за I_A – момент инерции корпуса относительно точки A , выраженный в подвижной системе координат; Ω_A – угловая скорость корпуса в той же системе координат.

Кинетическая энергия поступательного движения квадрокоптера задается по формуле:

$$T_{trans} = \frac{m}{2} (\dot{\xi}, \dot{\xi}).$$

Кинетическая энергия вращения корпуса:

$$T_{rotA} = \frac{1}{2} (\Omega_A, I_A \Omega_A),$$

где $I_A = \text{diag}(I_1^A, I_2^A, I_3^A)$, причем в силу симметрии $I_1^A = I_2^A$.

В силу Предположения 2 угловая скорость лопастей в подвижной системе координат выражается по формуле:

$$\Omega_{C_i} = \Omega_A + (-1)^{i+1} \dot{\gamma}_i e_3.$$

Предположим, что несущие винты в процессе собственного вращения образуют диск с центром масс в точке C_i и диагональным тензором инерции I_C в подвижной системе координат: $I_C = \text{diag}(I_1^C, I_2^C, I_3^C)$, причем в силу симметрии $I_1^C = I_2^C$.

Таким образом, суммарная кинетическая энергия квадрокоптера выглядит следующим образом:

$$\begin{aligned}
T_{rotA} &= \frac{1}{2}(\Omega_A; I_A \Omega_A) + \frac{1}{2} \sum_{i=1}^4 (\Omega_{C_i}; I_{C_i} \Omega_{C_i}) = \\
&= \frac{1}{2}(\Omega_A; (I_A + 4I_C) \Omega_A) + \frac{1}{2} I_3^c (\dot{\gamma}_1^2 + \dot{\gamma}_2^2 + \dot{\gamma}_3^2 + \dot{\gamma}_4^2) + \\
&+ (\dot{\gamma}_1 - \dot{\gamma}_2 + \dot{\gamma}_3 - \dot{\gamma}_4)(\Omega_A; I_C e_3).
\end{aligned}$$

Единственная потенциальная энергия, которая будет учтена, – энергия силы тяжести: $U = mgz$.

Угловая скорость в подвижной системе координат имеет следующий вид:

$$\Omega_A = \begin{pmatrix} \dot{\theta} - \dot{\phi} \sin \psi \\ \dot{\psi} \cos \theta + \dot{\phi} \cos \psi \sin \theta \\ \dot{\phi} \cos \psi \cos \theta - \dot{\psi} \sin \theta \end{pmatrix}$$

Положим по определению:

$$W_v := \begin{pmatrix} -\sin \psi & 0 & 1 \\ \cos \psi \sin \theta & \cos \theta & 0 \\ \cos \psi \cos \theta & -\sin \theta & 0 \end{pmatrix}$$

Тогда:

$$\dot{v} := \begin{pmatrix} \dot{\phi} \\ \dot{\psi} \\ \dot{\theta} \end{pmatrix} = W_v^{-1} \begin{pmatrix} \dot{\theta} - \dot{\phi} \sin \psi \\ \dot{\psi} \cos \theta + \dot{\phi} \cos \psi \sin \theta \\ \dot{\phi} \cos \psi \cos \theta - \dot{\psi} \sin \theta \end{pmatrix}$$

Заметим, что $\det W_v = -\cos \psi$, то есть кинематическое преобразование будет невырожденным для всех ориентаций за исключением тех, у которых $\psi = \pi/2 + \pi k$.

Полная функция Лагранжа имеет вид:

$$L(q, \dot{q}) = T_{trans} + T_{rot} - U.$$

1.3. Обобщенные силы

В соответствии с Предположением 3 существует четыре силы поступательного движения, и все они направлены вдоль оси e_3 . Следовательно, направление силы поступательного движения, приложенной к корпусу, определяется его ориентацией. В инерциальной системе координат направление силы тяги задается вектором: $G(\eta) = R e_3$.

Поэтому суммарную силу тяги, рассматриваемую в качестве управления, можно выразить в инерциальной системе отсчета следующим образом:

$$F_{\xi} = (u_1 + u_2 + u_3 + u_4) R e_3 = \begin{pmatrix} (u_1 + u_2 + u_3 + u_4)(\cos \varphi \sin \psi \cos \theta + \sin \varphi \sin \theta) \\ (u_1 + u_2 + u_3 + u_4)(\sin \varphi \sin \psi \cos \theta - \cos \varphi \sin \theta) \\ (u_1 + u_2 + u_3 + u_4) \cos \varphi \cos \theta \end{pmatrix}$$

Три вращательных момента из Предположения 4, являющиеся управлениями, имеют следующие выражения в подвижной системе координат:

$$\begin{aligned} \Gamma_1 &= (u_2 - u_4) l e_1, \\ \Gamma_2 &= (u_3 - u_1) l e_2, \\ \Gamma_3 &= (M_1 + M_2 + M_3 + M_4) e_3, \end{aligned}$$

где M_i – момент сопротивления i -того винта.

В итоге обобщенные силы, зависящие от переменных v , имеют вид:

$$\tau^0 := \begin{pmatrix} \tau_{\varphi} \\ \tau_{\psi} \\ \tau_{\theta} \end{pmatrix} = W_v^{-1} \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \end{pmatrix} = \begin{pmatrix} \frac{\sin \theta}{\cos \psi} (u_3 - u_1) l + \frac{\cos \theta}{\cos \psi} M \\ \cos \theta (u_3 - u_1) l + \sin \theta M \\ (u_2 - u_4) l + tg \psi (\sin \theta (u_3 - u_1) + \cos \theta M) \end{pmatrix}$$

1.4. Модель подъемной силы и момента несущего винта

В основе составления модели подъемной силы и момента несущего винта была положена статья [Avila Vilchis, 2003], в которой изучен вертолет, установленный на экспериментальной платформе. Рассматривая малый элемент лопасти, имеем следующие выражения для аэродинамической силы и момента:

$$\begin{aligned} u_i &= \frac{\rho p c a V R^2}{4} \dot{\gamma}_i V, \\ M_i &= \frac{\rho p c a R^4 c_d \dot{\gamma}_i^2}{8}, \end{aligned}$$

где ρ – плотность воздуха, p – количество винтов, c – ширина лопасти, a – угловой коэффициент кривой подъема, R – радиус диска

винта, V – индуцированная скорость парения, c_d – коэффициент сопротивления.

Введем дополнительные обозначения:

$$k_1 = \frac{\rho p c_a V R^2}{4}, k_2 = \frac{\rho p c_{c_d} V R^4}{8},$$

тогда:

$$u_i = (-1)^{i+1} k_1 \dot{\gamma}_i \text{ и } M_i = (-1)^i k_2 \dot{\gamma}_i^2.$$

1.5. Уравнения Лагранжа

Для начала введем обозначения: $A = I_1^a + 4I_1^c$, $D = I_3^a + 4I_3^c$.

Уравнения Лагранжа:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = Q,$$

где Q – обобщенные силы,

$$\begin{aligned} L(q, \dot{q}) &= T_{trans} + T_{rot} - U = \\ &= \frac{1}{2} A (\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \psi - 2\dot{\phi}\dot{\theta} \sin \psi + \dot{\psi}^2 \cos^2 \theta + \dot{\phi}^2 \cos^2 \psi \sin^2 \theta + 2\dot{\phi}\dot{\psi} \cos \theta \sin \theta \cos \psi) + \\ &+ \frac{1}{2} D (\dot{\psi}^2 \sin^2 \theta + \dot{\phi}^2 \cos^2 \psi \cos^2 \theta - 2\dot{\phi}\dot{\psi} \cos \theta \sin \theta \cos \psi) + \frac{1}{2} I_2^c (\dot{\gamma}_1^2 + \dot{\gamma}_2^2 + \dot{\gamma}_3^2 + \dot{\gamma}_4^2) \\ &- I_3^c (\dot{\gamma}_1 - \dot{\gamma}_2 + \dot{\gamma}_3 - \dot{\gamma}_4) (\dot{\phi} \cos \psi \cos \theta - \dot{\psi} \sin \theta). \end{aligned}$$

Стоит отметить, что полученную систему 6-ти дифференциальных уравнений второго порядка можно представить в виде $K \ddot{q} = M$, где K – симметричная матрица с определителем $A^2 D \cos 2\psi$, M – правая часть уравнений.

2. Моделирование и алгоритмы управления

Численное интегрирование полученной системы уравнений осуществлялось с помощью программы MATLAB R2013a. Был рассмотрен квадрокоптер массой 1 кг, с геометрическими параметрами $l = 0,4$ м, длина лопасти – 0,05 м, ширина – 0,01 м.

2.1. Взлет-посадка

В качестве одной из базовых траекторий рассмотрим вертикальный взлёт и посадку, в этом случае угловые скорости всех четырех винтов совпадают и равны ω . Выберем управление таким образом, чтобы $z(t)$ и $\omega(t)$ были гладкими непрерывными функциями. Результирующее решение приведено на рис. 2.

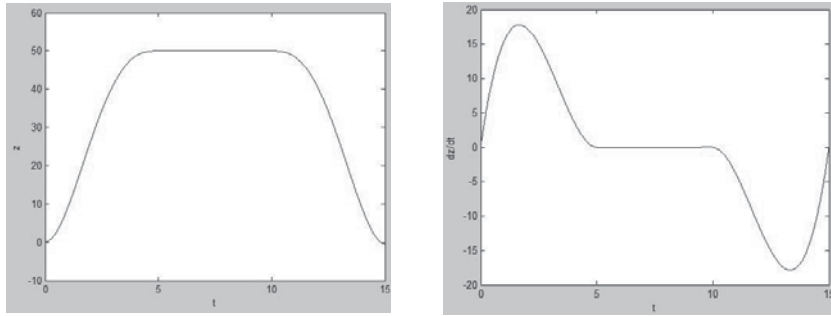


Рис. 2. График зависимости высоты и абсолютной скорости от времени при взлете-посадке

Для его реализации управление разделено на 3 этапа:

- 1) при $0 \leq t < 5$ $\omega(t)$ меняется по квадратичному закону, происходит набор высоты;
- 2) при $5 \leq t < 10$ $\omega = gm / 4k_1 = 37.73$ рад/с квадрокоптер зависает на высоте $h = 50$ м;
- 3) при $10 \leq t < 15$ $\omega(t)$ меняется по аналогичному квадратичному закону, происходит посадка.

Отметим, что на интервале $0 \leq t < 5$ координата $z(t)$ являлась функцией четвертой степени от t .

2.2. Полет по горизонтальной прямой

Одним из вариантов осуществления полета по прямой является движение квадрокоптера в плоскости Ozy при некотором фиксированном угле крена θ . Был рассмотрен случай $\theta = -0.216935$. Заметим, что при этом при дальности полета порядка 700 м высота изменилась всего на 2 мм (в идеальном смоделированном случае).

3. Нейросетевой контроллер

В задачах для систем с дефицитом управлений, некоторые из которых рассмотрены в книге [Фантони, 2013], управление строится исходя из свойств пассивности системы с использованием энергетического метода. В данном разделе предложим нейросетевой метод синтеза управления. Он, в частности, позволяет преодолеть проблему указанного дефицита.

3.1. Алгоритм синтеза управления

На основе результатов численного моделирования формируется обучающая выборка, которая покрывает пространство возможных движений. Далее в этом пространстве сеть проходит обучение, вычисляются необходимые весовые коэффициенты, после чего сеть используется для синтеза управлений. Рассмотрены 2 основных варианта. В первом использовались линейные датчики положений и углов аппарата (по данным навигационной аппаратуры), во втором – датчики ускорений по шести осям аппарата. В этом случае реализована композитная (гибридная) сеть, изображенная на рис. 3. Первые пары блоков 1-6 этой сети образуют рекуррентную сеть типа сети Джордана для вычисления положений по ускорениям (фактически, для двойного интегрирования ускорений), завершающий 7-й блок, это трехслойная сеть для окончательного расчета управлений.

Проведенное моделирование показало достаточно эффективную работоспособность этой схемы.

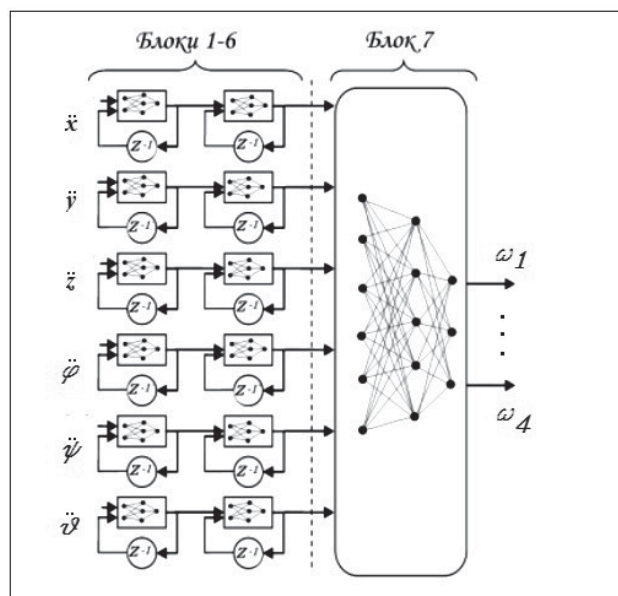


Рис. 3. Композитная нейросеть для синтеза управлений по датчикам ускорений

Заключение

Построена модель квадрокоптера с четырьмя управлениями $\dot{\gamma}_1, \dot{\gamma}_2, \dot{\gamma}_3, \dot{\gamma}_4$ и шестью степенями свободы $x, y, z, \varphi, \psi, \theta$, учитывающая сопротивление воздуха через воздействие его на вращение лопастей. С помощью численного моделирования изучены такие базовые траектории, как взлет-посадка и полет по прямой. Предложен метод синтеза управления на основе нейросетевого регулятора.

СПИСОК ЛИТЕРАТУРЫ

[Фантони, 2013] И. Фантони, Р. Лозано. Нелинейное управление механическими системами с дефицитом управляющих воздействий. – Москва – Ижевск: ООО «Компьютерная динамика», 2013, 312 с.

[Avila Vilchis, 2003] J.C. Avila Vilchis, B. Brogliato, A. Dzulf, R. Lozano, Nonlinear modelling and control of helicopters, Automatica 39, 2003.

[Bresciani, 2008] Tommaso Bresciani, Modelling, Identification and Control of a Quadrotor Helicopter, Department of Automatic Control, Lund University, October 2008.

[Hoffmann, 2007] G. M. Hoffmann, H. Huang, S. L. Waslander, and C. J. Tomlin, Quadrotor helicopter flight dynamics and control: Theory and experiment, Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit, Aug. 2007, 1–20.

[Prouty, 1995] Prouty R.W. Helicopter Performance, Stability, and Control. Krieger Publishing Company, 1995.

УДК 004.94:519.876.5

СРЕДА ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ГРУППОВЫХ ДЕЙСТВИЙ АВТОНОМНЫХ БПЛА

М.Г. Пантелеев (*MPanteleyev@gmail.com*),
С.В. Лебедев (*lebedev.sv.etu@gmail.com*)
СПбГЭТУ «ЛЭТИ», Санкт-Петербург
Н.В. Кохтенко (*nvkokhtenko@gmail.com*),
ЗАО НИИ ТС «Синвент», Санкт-Петербург

Рассмотрены проблемы создания среды имитационного моделирования (СИМ) групповых действий БПЛА с высокой степенью автономности. Среда предназначена для экспериментального исследования и отработки в реальном времени различных моделей группового поведения БПЛА в условиях командного противодействия и разрабатывается с позиций концепции многоагентных систем (МАС). Сформулированы предъявляемые к СИМ требования, рассмотрена ее архитектура и основные особенности функционирования.

Введение

Магистральной тенденцией развития БПЛА является повышение степени их автономности [Huang et al., 2008], в том числе при решении задач группами в условиях активного противодействия [Baxter et al., 2008; Stenger et al., 2012]. Теоретической базой создания таких систем является концепция автономных интеллектуальных агентов (ИА) и основанных на них многоагентных систем (МАС). Практическая разработка автономных БПЛА, способных координировать свои действия при групповом решении задач, предполагает создание технологических средств, поддерживающих их проектирование и исследование поведения. Натурное исследование моделей автономного и группового поведения БПЛА в различных ситуациях экономически нецелесообразно, поэтому актуальным является создание сред имитационного моделирования таких систем [Selecký et al., 2012]. Существующие в настоящее время среды имитационного моделирования агентных систем по совокупности характеристик не удовлетворяют требованиям моделирования групповых действий БПЛА. Таким образом, актуальной проблемой является создание платформы, позволяющей исследо-

вать групповые действия БПЛА в широком диапазоне сценариев с учетом требований реального времени.

1. Основные требования к СИМ

Основными требованиями к СИМ являются:

- *универсальность* – поддержка моделирования различных типов БПЛА как в части физических подсистем (ФП), так и архитектуры интеллектуальной подсистемы (ИП);
- *открытость* – возможность добавления новых компонентов, используемых при построении имитационных моделей БПЛА (например, новых типов сенсоров), включая его ИП;
- *масштабируемость* – возможность наращивания числа моделируемых объектов без снижения общей производительности системы.

Важнейшим требованием к СИМ является обеспечение возможности моделирования сценариев группового противодействия в реальном масштабе времени. При этом длительность такта моделирования должна соответствовать динамике процессов в моделируемой физической среде. Автономные БПЛА, функционирующие в открытых динамических мирах, с позиций теории ИА следует рассматривать как системы «ограниченной рациональности». Такие системы, в общем случае, не могут за отведенное время оптимально решать стоящие перед ними задачи, а должны рационально использовать имеющееся в конкретной ситуации время обдумывания решения. В высокодинамичных средах решения должны приниматься агентом в темпе обновления данных о состоянии среды. Вследствие этого ИП БПЛА должна выдавать команды управления физической подсистеме, а СИМ – обеспечивать обработку этих команд (т.е. пересчет физического состояния всех моделируемых БПЛА) в пределах такта моделирования.

Другое важное требование к создаваемой СИМ – возможность мониторинга работы ИП БПЛА (ментальной деятельности агента). Это предполагает фиксацию в каждом такте моделирования промежуточных результатов процесса обдумывания решений на каждом значимом для используемых архитектуры и модели ИП шаге принятия решений. Соответствующая информация должна сохраняться для последующего воспроизведения и анализа сеанса моделирова-

ния вне режима реального времени. Данное требование обусловлено необходимостью сравнительного исследования эффективности различных архитектур ИП как систем «ограниченной рациональности».

Еще одним требованием к СИМ является возможность моделирования различных стратегий координации действий в группах БПЛА. В настоящее время в теории МАС разработано множество подходов и моделей координации поведения агентов, ориентированных на различные классы агентных систем и решаемые ими задачи. Выбор конкретных стратегий и моделей координации выполняется разработчиком моделируемой МАС, при этом СИМ не должна накладывать в этой части никаких ограничений. Задачей среды является предоставление системных механизмов, позволяющих моделировать различные стратегии координации как при наличии непосредственной коммуникации между агентами, так и без нее. В первом случае СИМ должна поддерживать моделирование каналов связи между агентами с учетом их ограниченной пропускной способности и дальности действия, наличия шумов и ненадежной доставки сообщений. Координация в условиях отсутствия непосредственной коммуникации предполагает наблюдение за поведением других агентов посредством собственных сенсоров и (или) получение информации о других агентах от единого центра управления (выделенного агента-координатора).

2. Архитектура и особенности реализации СИМ

Архитектура СИМ, удовлетворяющей сформулированным выше требованиям, представлена на рис. 1. В ее составе выделены среда разработки и среда исполнения.

Среда разработки позволяет создавать новых агентов (БПЛА и другие, в том числе наземные, активные сущности), задавать начальные условия сценариев моделирования и конфигурировать среду исполнения. При создании ФП используются библиотеки классов, реализующих различные модели динамики физического носителя агента (например, управляемого полета беспилотного летательного аппарата).

Разрабатываемая СИМ ориентирована, в первую очередь, на исследование агентов с делиберативной архитектурой ИП. Такие агенты используют явную символьную модель окружающей среды

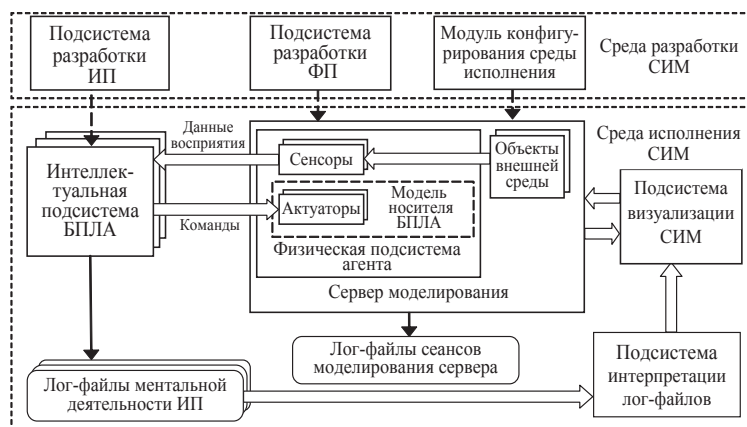


Рис. 1. Архитектура СИМ

и основанные на этой модели методы прогнозирования событий и планирования действий. Класс таких архитектур достаточно широк. В перспективе СИМ не должна накладывать ограничений на выбор конкретной архитектуры ИП. На данном этапе в качестве базовой архитектуры ИП выбрана архитектура, основанная на полезности [Пантелеев, 2004; Пантелеев, 2012], включающая подсистемы оценки обстановки, прогнозирования и планирования действий.

В соответствии с архитектурой выполняется разработка отдельных подсистем ИП. При построении разных подсистем используются различные модели представления и обработки знаний – продукционные, фреймовые, а также различные модели обработки неопределенности. Программная архитектура допускает использование существующих инструментов построения систем, основанных на знаниях (CLIPS, JESS и др.). Среда разработки ИП поддерживает редактирование баз знаний агентов и конфигурирование методов их обработки.

Среда исполнения обеспечивает реализацию сеансов моделирования и включает следующие основные компоненты: сервер моделирования, подсистему визуализации и подсистему интерпретации лог-файлов сеансов моделирования. Модуль конфигурирования среды исполнения позволяет задавать длительность такта моделирования, режим работы подсистемы визуализации и логгирования процесса моделирования и т.п.

Среда исполнения СИМ имеет клиент-серверную архитектуру. В соответствии с требованиями масштабирования и обеспечения реального времени ИП различных агентов реализуются в отдельных процессах и могут выполняться на отдельных компьютерах локальной сети. Подсистема визуализации также поддерживает клиент-серверное взаимодействие с сервером моделирования и может быть реализована на выделенном сервере для отображения моделируемых процессов в 3D-графике, что требует значительных вычислительных ресурсов.

Подсистема интерпретации лог-файлов предназначена для воспроизведения сеансов моделирования вне режима реального времени с возможностью детального анализа динамики ментальной деятельности агентов. Поскольку моделирование состояния многоагентного физического мира и ментальной деятельности отдельных агентов в общем случае выполняется на разных компьютерах, лог-файлы этих процессов также сохраняются на разных компьютерах. Состав сохраняемой информации в общем случае зависит от выбранной архитектуры агента и моделей принятия решений. В частности, для архитектуры ИА, основанной на полезности, эта информация может включать варианты действий, текущие значения их полезности, прогнозируемые ситуации и др. В режиме последующего воспроизведения сеанса моделирования подсистема интерпретации лог-файлов работает с распределенной файловой системой и обеспечивает отображение физического процесса и состояния ментальной деятельности агентов в каждом такте. Базовый алгоритм работы сервера моделирования представлен на рис. 2.

В каждом такте моделирования через заданный интервал времени Δt сервер опрашивает порты, к которым подключены ИП моделируемых агентов, и считывает поступившие от них команды. Такт работы сервера задает интервал, в рамках которого ИП должны принимать сенсорные данные, обрабатывать их, вырабатывать решения и выдавать команды управления. На основе этих команд сервер вычисляет новые физические состояния всех агентов (состояния ФП). При этом реализуемые модели агентов как систем ограниченной рациональности должны гарантировать выполнение цикла работы ИП в рамках заданного такта моделирования.

Новые зоны сенсорного покрытия определяются для всех агентов с учетом нового пространственного положения агентов и ре-

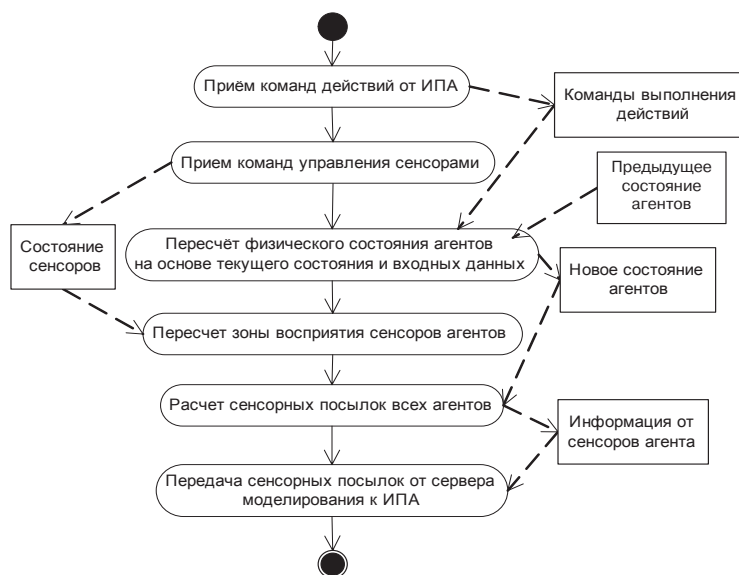


Рис. 2. Базовый алгоритм работы сервера моделирования

жимов работы их сенсоров. При этом для каждого агента определяются объекты внешней среды, находящиеся в пределах зоны его восприятия, и вычисляются сенсорные данные с учетом заданной модели ошибок. СИМ поддерживает возможность моделирования различных типов сенсоров, входящих в состав подсистемы восприятия (ПВсп) агента. Существенными особенностями имитационного моделирования сенсоров агентных систем являются:

- ограниченность зон восприятия;
- возможность управления режимами работы сенсора;
- наличие ошибок (шумов) в воспринимаемой информации.

Среда разработки СИМ поддерживает возможность создания новых типов сенсоров, а также использования при построении конкретного ИА существующих типов сенсоров с возможностью конфигурирования их параметров. При создании имитационной модели сенсора задаются параметры зоны восприятия, характеристики объектов, воспринимаемые данным типом сенсора, и модель ошибок измеряемых параметров. При наличии у сенсора разных режимов работы для каждого из них определяются собственные характеристики.

В общем случае модель внесения погрешностей в получаемые сенсором данные имеет вид:

$$\Delta = f(d, R, P),$$

где Δ – относительная погрешность, вносимая в измеряемые данные; d – расстояние до наблюдаемого объекта; R – режим работы сенсора; P – текущая характеристика помеховой обстановки; f – функция, описывающая зависимость величины погрешности от указанных выше параметров. При моделировании помеховой обстановки учитывается, что она может быть обусловлена как естественными шумами, так и помехами, целенаправленно создаваемыми другими агентами.

Система позволяет моделировать произвольное число агентов, подсистема восприятия каждого агента может содержать любое число сенсоров различных типов. С учетом этого алгоритм работы сервера, реализующий моделирование восприятия агентов, представлен на рис. 3.

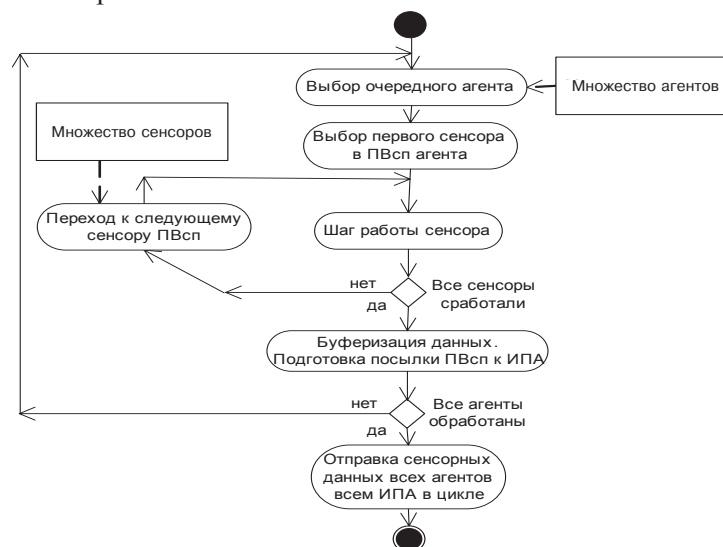


Рис. 3. Базовый алгоритм моделирования восприятия агентов

Собранные в текущем такте сенсорные данные сериализуются и отсылаются ИП соответствующих агентов. На основе полученных сообщений ИП обновляет внутреннюю модель мира агента

текущими значениями параметров наблюдаемых объектов, вычисляет значения высокоуровневых параметров ситуации и выполняет ее оценку. Обобщенный алгоритм работы сенсора выполняет обход графа сцены и выбор объектов, попавших в зону восприятия сенсора.

Для отработки моделей кооперативного поведения БПЛА, основанных на коммуникации, в СИМ реализована возможность моделирования ненадежных каналов связи с ограниченной пропускной способностью. Поддерживается возможность моделирования как широкополосных каналов связи, так и выделенных между парой агентов («точка–точка»). При создании модели канала связи в среде разработки задаются его характеристики, в частности: пропускная способность, вероятность безошибочной доставки сообщения, максимальная дальность связи (при моделировании радиоканалов) и др. Указанные характеристики могут описываться случайными величинами с задаваемыми параметрами распределения. Поддерживается возможность приема нескольких сообщений от различных агентов в одном такте моделирования.

Модели каналов связи реализуются основным сервером моделирования. Связь реализуется посылкой серверу команды, содержащей сообщение и идентификатор получателя. В следующем такте принимающий агент получает от сервера команду, содержащую переданное сообщение. При моделировании радиоканалов сервер добавляет во входное сообщение направление на источник и расстояние до него с внесением погрешностей относительно истинных значений. Таким образом, поддерживается возможность с определенной степенью точности идентифицировать источник сообщения по его положению в пространстве.

Заключение

Разрабатываемая среда имитационного моделирования позволяет исследовать модели автономного и группового поведения БПЛА в широком диапазоне сценариев с учетом реальных ограничений их сенсоров и каналов связи. СИМ поддерживает режим реального времени и возможность журналирования и анализа работы ИП (ментальной деятельности агентов) в процессе принятия решений в динамических мирах, что позволяет обрабатывать различные модели принятия решений.

СПИСОК ЛИТЕРАТУРЫ

[**Huang et al., 2008**] Huang H., Albus J, Messina E. Specifying Autonomy Levels for Unmanned Systems: Interim Report //SPIE Defense and Security Symposium, Orlando, Florida, 2004.

[**Baxter et al., 2008**] Baxter J. W., Horn G. S., Leivers D. P. Fly-by-agent: Controlling a pool of UAVs via a multi-agent system. *Knowledge-Based Systems*, **21**(3): 232–237, 2008.

[**Stenger et al., 2012**] Stenger A., Fernando B., Heni M. Autonomous Mission Planning for UAVs: A Cognitive Approach// Deutscher Luft- und Raumfahrtkongress 2012, DocumentID: 281398. <http://www.dglr.de/publikationen/2013/281398.pdf>.

[**Selecký et al., 2012**] Selecký M., Meiser T. Integration of Autonomous UAVs into Multi-agent Simulation// Acta Polytechnica Vol. 52 No. 5/2012, pp. 93–99.

[**Пузанков и др., 2008**] Пузанков Д.В., Мирошников В.И., Пантелеев М.Г., Серегин А.В. Интеллектуальные агенты, многоагентные системы и семантический Web: концепции, технологии, приложения. – СПб.: ООО «Технолит», Изд-во «Технолит», 2008.

[**Пантелеев, 2004**] Пантелеев М.Г. Планирование действий интеллектуального агента в реальном времени// Труды 9-й национальной конф. по искусственному интеллекту с международным участием КИИ-2004. – М.: Физматлит, 2004. Т 2.

[**Пантелеев, 2012**] Пантелеев М.Г. Концепция построения интеллектуальных агентов реального времени на основе модели опережающего итеративного планирования// Труды 13-й национальной конф. по искусственному интеллекту с международным участием КИИ-2012. – Белгород: Изд-во БГТУ, 2012. Т 3.

УДК 004.023, 004.921, 004.932.72'1

**ПРОГРАММНАЯ СИСТЕМА ДЛЯ МОДЕЛИРОВАНИЯ
ПОЛЕТА БПЛА И ЕГО АВТОНОМНОГО
ПОЗИЦИОНИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ
И АЛГОРИТМОВ ТЕХНИЧЕСКОГО ЗРЕНИЯ**

Д.Н. Степанов (*mitek1989@mail.ru*)

И.П. Тищенко (*billy@billy.botik.ru*)

Институт программных систем им. А.К. Айламазяна РАН,
Переславль-Залесский

В работе описывается программная система, предназначенная для моделирования полета беспилотного летательного аппарата с помощью средств, методов и алгоритмов технического зрения. Представлены общая схема генерации и обработки данных, а также краткое описание модулей программной системы, предназначенные для решения отдельных подзадач.

Введение

Задачи, которые ставятся перед БПЛА, обусловлены их многочисленными приложениями в военной и гражданской сферах (картографирование, сельское хозяйство, охранные мероприятия, мониторинг и т. д.). Создание надежных систем навигации и управления БПЛА – важное направление разработок. Стабильность движения большинства БПЛА существенно зависит от качества сигналов спутниковой связи GPS/ГЛОНАСС. Общий недостаток любой радионавигационной системы – при определенных условиях сигнал может не доходить до приемника, или приходиться со значительными искажениями или задержками (городские постройки, рельеф местности, облачность). Кроме того, особенность GPS – зависимость условий получения сигнала от Минобороны США. Инерциальные навигационные микромеханические системы (ИНМС) способны работать автономно, но их использование неизбежно связано с проблемой накопления ошибки в вычислениях, что с течением времени приводит к большим погрешностям.

Поэтому создание системы технического зрения, которая позволила бы меньше зависеть от спутниковой навигации и ИНМС, является актуальной, перспективной и важной научной задачей, и

востребована разработчиками и производителями БПЛА. Имеется ряд эффективных методов компьютерной графики, искусственного интеллекта и параллельных вычислений, которые могут быть использованы в качестве интеллектуальной основы при разработке методов и алгоритмов решения задачи автономного позиционирования с применением фото- или видеокамер.

Коллектив исследователей ИПС РАН участвовал в научно-исследовательской работе «Разработка алгоритмического, математического и программного обеспечения системы моделирования автономного позиционирования БПЛА (аппаратно-программного комплекса) на основе технологий машинного зрения, искусственного интеллекта и параллельных вычислений». В рамках проекта был создан прототип программной системы, предназначенной для моделирования полета БПЛА и его автоматического позиционирования при полете над плоской местностью. В настоящий момент разрабатывается и реализуется система моделирования полета аппарата и позиционирования при полете аппарата над местностью с рельефом.

1. Постановка задачи и общее описание программной системы

Постановка задачи заключается в следующем: задан спутниковый снимок местности, над которой БПЛА, оснащенный фото- или видеокамерой, совершает свои полеты. Для кадров с камеры на БПЛА (она направлена вниз) необходимо определить, где в эти моменты находился аппарат и, возможно, определить ориентацию аппарата относительно Земли.

В основе решения задачи лежит использование искусственного спутника Земли (ИСЗ) для тех участков местности, над которой БПЛА совершает свои полеты. Снимок имеет геопривязку. На любом снимке с БПЛА будет запечатлен некоторый фрагмент снимка со спутника. Вычисляются параметры проективного преобразования, которое связывает координаты ряда точек, наблюдаемых на обоих снимках, что позволяет рассчитать положение и ориентацию камеры на БПЛА сначала относительно положения камеры на спутнике, а затем в глобальной системе координат.

Эти же спутниковые снимки используются для моделирования полета. Если необходимо учитывать рельеф местности и кривизну

поверхности Земли, в дополнение к спутниковому снимку используется карта высот. Это псевдоизображение, в котором яркость каждого пикселя – высота соответствующей точки над уровнем моря.

Программная система функционирует на кластерной вычислительной (КВУ), которая состоит из управляющей машины (фронтэнда) и вычислительных узлов, КВУ находится под управлением одного из вариантов ОС Linux. Главные особенности системы: обработка данных в так называемом конвейерно-параллельном режиме и модульная архитектура [Тищенко и др., 2012]. Вычислительное ядро отвечает за решение задач планирования и распределения нагрузки, основная же функциональность заложена в модулях, каждый из которых может иметь произвольное назначение. Модульность позволяет легко расширять функциональность системы для решения самых разных задач, главное требование: модули, образующие логическую схему обработки, должны подходить друг другу по типу входных и выходных данных.

Рассмотрим разработанные программные модули, которые функционируют в рамках программной системы и предназначены для решения отдельных подзадач.

2. Подготовка исходных данных

Эта задача решается один раз, если у нас зафиксирован спутниковый снимок, который используется для моделирования полета и позиционирования. Для спутникового снимка строится так называемая пирамида изображений: в основании лежит исходный снимок, а каждый следующий слой по ширине и высоте в 2 раза меньше предыдущего. Все изображения нарезаются на квадратные фрагменты одинакового размера. На каждом фрагменте с помощью алгоритма SURF [Bay et al., 2008] находятся так называемые особые точки, для каждой точки вычисляется ее дескриптор (набор вещественных чисел). Причина разбиения снимка на фрагменты и построение пирамиды изображений будут обоснованы далее.

3. Генерация траектории

Создание реалистичных траекторий полета не является первоочередной задачей исследований, поэтому было решено поступить следующим образом: пользователь отмечает на уменьшенной ко-

пии спутникового снимка ряд опорных точек, которые соединяются прямыми отрезками. На полученной цепной линии вычисляются координаты ряда точек A_i (соседние точки равноудалены друг от друга). Этим точкам будут соответствовать те моменты времени, в которые камера на виртуальном БПЛА выполняет съемку. Также для каждой точки вычисляется ориентация «беспилотника» в глобальной системе координат (навигационные углы: тангаж, крен и рысканье). Если учитывать кривизну поверхности Земли, то эти отрезки будут прямыми только в криволинейной системе координат, которая определяется поверхностью земного эллипсоида. В геоцентрических (прямоугольных) координатах это будут кривые линии с достаточно сложными уравнениями.

Решение задачи можно разделить на два этапа. Первый – вычисление расстояния между двумя соседними опорными точками (вычисление длины кривой линии). Математически задача сводится к вычислению определенного интеграла от функции, первообразная которой не выражается в элементарных функциях, и поэтому расстояние проще найти приближенно, разделив кривую линию на множество прямых отрезков (т.е. используя аналог одного из численных методов вычисления определенного интеграла). На втором этапе кривая линия делится на некоторое количество частей, зависящее от ее длины и расстояния, которое пролетает БПЛА между моментами съемки двух соседних кадров видеоряда. Затем вычисляются координаты всех промежуточных точек.

4. Порождение видеоряда

Каждой точке A_i сопоставляется свой кадр видеоряда. Задача сводится к перспективному преобразованию исходного спутникового снимка с помощью так называемой матрицы гомографии, которая вычисляется исходя из положения и ориентации камеры на БПЛА, а также параметров камер на ИСЗ и БПЛА. При этом виртуальная камера располагается так, что при нулевом угле тангажа БПЛА камера направлена перпендикулярно поверхности земного эллипсоида.

Учет рельефа местности и кривизны Земли обуславливает применение более продвинутых средств: систем трехмерного моделирования (например, *osgEarth*), предназначенных для визуализации 3D-моделей. Пользователь с помощью мыши или клавиатуры по-

зиционирует виртуальную камеру, которая «парит» над поверхностью Земли (пример показан на рис. 1). В нашем случае нет необходимости отображать сгенерированные снимки, более предпочтительно проводить так называемый «offscreen rendering» – процесс, в результате которого сгенерированное изображение не выводится на экран компьютера, а остается в оперативной памяти.

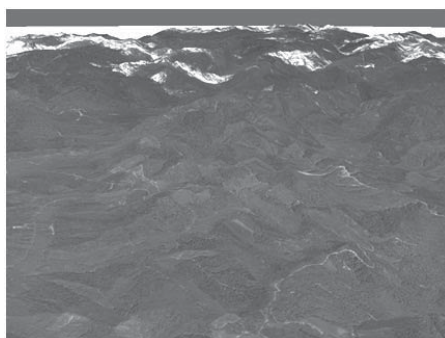


Рис. 1. Пример сгенерированного кадра видеоряда

Чаще всего пользователи работают с КВУ в удаленном режиме посредством клиентских ПК и сетевых протоколов взаимодействия. Компьютер клиента служит только для ввода/вывода информации, а все вычисления проводятся на сервере. Серия экспериментов с osgEarth показала, что обеспечить корректный рендеринг трехмерных сцен в удаленном режиме удается не всегда. Это связано с самим принципом работы графических приложений и недостаточной развитостью программных средств, которые предназначены для отображения интерфейсов программ, работающих в удаленном режиме. Вторая причина состоит в том, что узлы кластеров если и оснащаются видеокартами, то специализирующимися именно на вычислениях, а не на отображении графики (например, серия видеокарт NVIDIA Tesla).

Поэтому предложен следующий подход: рендеринг трехмерных сцен происходит на пользовательской машине, а КВУ производит все остальные вычисления. КВУ посылает на ПК набор чисел, которые задают положение и ориентацию камеры в пространстве. Компьютер пользователя производит рендеринг сцены и отправляет сгенерированное изображение на КВУ. Связь КВУ и ПК можно

осуществить с применением стандартных TCP-сокетов. При таком подходе КВУ выступает в качестве клиента, а ПК – в качестве сервера. Копии одного и того же спутникового снимка хранятся и там, и там. Пример сгенерированного кадра видеоряда представлен на рис. 1.

5. Поиск пар общих точек на снимках с БПЛА и ИСЗ

Задан кадр видеоряда, на нем с помощью алгоритма SURF находятся особые точки и вычисляются их дескрипторы. Далее они сопоставляются с особыми точками со спутникового снимка (в качестве метрики используется евклидово расстояние между векторами-дескрипторами). Поскольку для предыдущего кадра видеоряда уже вычислены положение и ориентация в глобальной системе координат, это позволяет выполнять сопоставление с особыми точками не со всего спутникового снимка (размер которого – от нескольких сотен до нескольких десятков тысяч мегапикселей), а только с некоторого подмножества его фрагментов с определенного слоя пирамиды, который соответствует некоторой ориентировочной высоте полета БПЛА. Для выбора нужных фрагментов используется разработанный алгоритм сужения области поиска [Степанов, 2014].

6. Поиск проективного преобразования, связывающего снимки с БПЛА и ИСЗ

Имеется множество пар координат общих особых точек со снимков с БПЛА и ИСЗ. В случае моделирования полета над плоской местностью и без учета кривизны Земли оба снимка являются проективными преобразованиями плоскости, и поэтому координаты пар особых точек связаны матрицей гомографии H . Для вычисления матрицы H необходимо как минимум 4 пары точек, использование большего количества точек позволяет повысить устойчивость вычислений (с применением метода наименьших квадратов). Но поскольку алгоритм сопоставления точек также основан на эвристиках, это может стать причиной зашумленности данных. Существуют стабильные методы оценки параметров некоторой модели на основе случайных выборок, позволяющие уменьшить влияние шума, один из них носит название RANSAC [Zuliani, 2008]. Далее производится декомпозиция матрицы H на матрицу

поворота (она определяет ориентацию камеры на БПЛА в системе координат, привязанных к ИСЗ) и вектор параллельного переноса (он определяет положение). Зная положение и ориентацию камеры на ИСЗ в глобальной системе координат, легко вычислить, где находится и как ориентирована камера на БПЛА.

В общем случае, когда на двух снимках запечатлена одна и та же сцена, не являющаяся плоской, координаты точек на снимках одной и той же сцены с двух разных камер в разных ракурсах связаны так называемой фундаментальной матрицей F . Существуют различные алгоритмы вычисления матрицы F , обзор показал, что наиболее устойчивым является пятиточечный алгоритм. Кроме того, он корректно работает в той ситуации, когда все пять выбранных пар точек располагаются на одной плоскости, другие алгоритмы в такой ситуации выдают некорректные результаты. Как и в случае с матрицей гомографии, для уменьшения влияния шума целесообразно использовать данный алгоритм в паре с одним из стабильных методов оценки параметров некоторой модели на основе случайных выборок (например, RANSAC).

Фундаментальную матрицу также можно разложить на матрицу поворота и вектор параллельного переноса (они определяют взаимное положение и ориентацию камер). Но вектор будет определен только с точностью до некоторого положительного коэффициента k . Для вычисления k необходимо знать трехмерные координаты хотя бы одной точки, образ которой есть на снимке с БПЛА. Координаты этой точки на спутниковом снимке известны. Для повышения устойчивости можно использовать сразу несколько таких точек, применяя связку «метод наименьших квадратов + RANSAC» для вычисления k .

В общем случае, проекция каждой такой точки может располагаться на всем снимке с БПЛА, но наличие фундаментальной матрицы позволяет использовать так называемые эпиполярные ограничения: поиск образов точек на снимке с БПЛА следует производить вдоль эпиполярных линий, уравнение которых зависит от матрицы F . Поскольку матрица F также может быть вычислена с некоторой погрешностью, предлагается рассматривать те пиксели снимка с БПЛА, которые располагаются на эпиполярной линии и на расстоянии в несколько пикселей от нее.

Далее встает вопрос: если известны координаты образа некоторой точки на снимке со спутника, то как найти образ этой точ-

ки на снимке с БПЛА? Необходимо выбрать некоторую метрику, которая позволяет сравнивать точки с двух изображений. Анализ источников показал, что одним из наиболее эффективных методов являются вейвлеты (фильтры) Габора [Kamagainen et al., 2006]. Рассматривается область некоторого радиуса вокруг точки, производится цифровая свертка этой области с 40 различными масками Габора, в итоге имеем 40-мерный вектор вещественных чисел – дескриптор точки. В качестве метрики для сравнения двух дескрипторов предлагается использовать косинус угла между векторами: чем он ближе к единице, тем больше точки с двух изображений «похожи» друг на друга. Фильтры Габора обладают свойством инвариантности к аффинным преобразованиям изображения и изменениям яркости.

На этапе подготовки исходных данных на каждом из фрагментов из каждого слоя пирамиды спутниковых снимков выбирается один случайный пиксель, с помощью семейства фильтров Габора вычисляется 40-мерный дескриптор пикселя, а наличие карты высот позволяет вычислить трехмерные координаты точки, чей образ находится в этом пикселе. Далее вся информация сохраняется для дальнейшего использования.

7. Использование GPU-вычислений в рамках программной системы

В последние несколько лет все более актуальным становится перенос различных алгоритмов на графические процессорные устройства, GPU позволяют добиться существенного ускорения в задачах с так называемым мелкогранулярным массовым параллелизмом (иногда в десятки и сотни раз), хотя и имеют свою специфику написания программ для них. Существуют различные программно-аппаратные платформы параллельных GPU-вычислений, наиболее популярной является CUDA.

В разрабатываемой программной системе есть несколько алгоритмов, для которых возможны эффективные GPU-аналоги. Например, для алгоритма выделения особых точек SURF существует несколько реализаций на CUDA. Другой алгоритм – поиск образа точки на снимке с БПЛА с известными ее координатами на снимке с ИСЗ. Поиск затрагивает эпиполярную линию и все близлежащие пиксели. Каждая вычислительная GPU-нить обрабатывает свой пиксель, принадлежащий снимку с БПЛА, и в качестве результа-

та выдает «расстояние» (по описанной выше метрике) между двумя дескрипторами, один из которых описывает пиксель с БПЛА-снимка, а второй – пиксель с ИСЗ-снимка.

Заключение

Разработаны и реализованы модули программной системы, отвечающие за моделирование полета виртуального БПЛА и позиционирования при полете аппарата над местностью с рельефом. На GPU реализован алгоритм поиска образа некоторой точки на снимке с БПЛА при известных ее координатах на снимке с ИСЗ. Технология обработки данных включает в себя удаленный персональный компьютер, который используется для порождения видеоряда с борта виртуального БПЛА.

В ближайшее время планируется провести всестороннее тестирование как отдельных модулей, так и всей схемы обработки данных (моделирование и позиционирование). В случае положительных результатов полученные наработки могут быть применимы для реализации программно-аппаратных комплексов, предназначенных для решения задач автономного позиционирования БПЛА, а также других подвижных технических средств.

СПИСОК ЛИТЕРАТУРЫ

[**Степанов, 2014**] Степанов Д.Н. Методы и алгоритмы определения положения и ориентации беспилотного летательного аппарата с применением бортовых видеокамер // Программные продукты и системы (международный журнал): 2014, т. 1, № 1. – <http://www.swsys.ru/index.php?page=article&id=3776>.

[**Тищенко и др., 2012**] Тищенко И.П., Степанов Д.Н., Фраленко В.П. Разработка системы моделирования автономного полета беспилотного летательного аппарата // Программные системы: теория и приложения. – электрон. научн. журн., 2012, т. 3, № 3(12). – http://psta.psiras.ru/read/psta2012_3_3-21.pdf.

[**Bay et al., 2008**] Bay H., Ess A., Tuytelaars T., Van Gool L. SURF: Speeded Up Robust Features // Computer Vision and Image Understanding (CVIU): 2008, Vol. 110, № 3.

[**Kamarainen et al., 2006**] Kamarainen J.-K., Kyrki V., Kälviäinen H. Invariance Properties of Gabor Filter Based Features – Overview and Applications // IEEE Transactions on Image Processing: Vol. 15, No. 5, 2006.

[**Zuliani, 2008**] Zuliani M. RANSAC for Dummies // 2008.

УДК 004.8

ЗАДАЧА ОБЛЕТА ТЕРРИТОРИИ БЕСПИЛОТНЫМ ЛЕТАТЕЛЬНЫМ АППАРАТОМ С УЧЕТОМ ВЕТРОВЫХ НАГРУЗОК¹

М.В. Хачумов (*khmike@inbox.ru*)
Институт системного анализа РАН, Москва

В работе решается задача облета беспилотным летательным аппаратом территории, заданной опорными точками, в условиях ветровой нагрузки и ограничения на управление. Она разбивается на задачу планирования траектории облета и задачу следования по выбранному маршруту. В соответствии с принципом раздельного движения рассматривается плоский случай. Решение задачи следования основывается на производственных правилах.

Введение

Возможность планирования оптимальной траектории полета по маршруту в реальном времени с учетом расстояния между опорными точками и воздействий на летательный аппарат (ЛА) способствует снижению эксплуатационных затрат за счет выбора наиболее экономичных маршрутов. В работе [Хачумов, 2010] задача планирования траектории движения летательного аппарата решается на основе венгерского метода [Кудрявцев, 1984] путем введения ограничений на перемещения и учета порядка облета пунктов. В настоящей работе планирование траектории полета с возвратом в исходную точку рассматривается как решение задачи коммивояжера на основе искусственной нейронной сети Кохонена. После того как маршрут облета территории рассчитан, необходимо решить задачу следования по нему с минимальными временными и координатными отклонениями. Моделирование следования БПЛА по заданному маршруту решается как задача преследования цели. Ранее подобный подход был предложен и рассмотрен автором в работе [Абрамов и др., 2013].

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 13-07-00025 А и № 14-07-31020 мол_а) и проекта Программы фундаментальных исследований президиума РАН № 18.

1. Постановка задачи

Пусть заданы n опорных точек (x_i, y_i) , $i = 1, \dots, n$ и известны расстояния между ними $d_{ij} = d((x_i, y_i), (x_j, y_j))$. Задача коммивояжера заключается в отыскании самого короткого маршрута, проходящего через каждую из n точек по одному разу с последующим возвратом в исходную позицию [Кузин, 1973]. Задача относится к классу NP – полных задач, где число вариантов оценивается как $N!$

Если ввести переменные:

$$x_{ij} = \begin{cases} 1, & \text{если БПЛА из точки } i \text{ движется в точку } j; \\ 0, & \text{в противном случае,} \end{cases}$$

то задача формулируется следующим образом:

$$\sum_{i=0}^n \sum_{j=0}^n d_{ij} x_{ij} \rightarrow \min; \quad \sum_{i=0}^n x_{ij} = 1, \forall j; \quad \sum_{j=0}^n x_{ij} = 1, \forall i, \text{ где:} \quad (1.1)$$

$$\begin{cases} j = 1, \dots, n; \quad i = 1, \dots, n; \\ u_i - u_j + nx_{ij} \leq n-1, \quad i \neq j. \end{cases} \quad (1.2)$$

Требования цикличности записываются в виде (1.2), где u_i – номер шага, на котором посетили точку i .

Решение задачи требует относительно больших затрат времени, что не позволительно при решении оперативно-тактических задач с помощью БПЛА. Известны эвристические методы решения задачи коммивояжера на основе ИНС Кохонена и Хопфилда [Ежов и др., 1998]. Заметим, что решение задачи коммивояжера на нейронной сети Хопфилда требует введения ряда параметров, которые можно определить только экспериментально путем подбора, что невозможно в условиях оперативных действий. В отличие от этого ИНС Кохонена позволяет получать достаточно быстро приемлемое решение для задачи большой размерности в условиях дефицита времени, поэтому выбор между указанными ИНС сделан в пользу последней.

Недостатком постановки (1.1)-(1.2) является отсутствие учета ветровых нагрузок на летательный аппарат, которые носят случайный характер и могут вызывать существенные отклонения от маршрута. Предлагается на первом этапе решать в реальном време-

ни задачу выбора маршрута, а на втором – задачу его облета уже с учетом действующих случайных нагрузок и имеющейся математической модели летательного аппарата. В соответствии с принципом раздельного движения рассматривается плоский случай решения задачи следования по маршруту с управлением углом тангажа θ в установленных моделью диапазонах и двухпозиционным управлением скоростью летательного аппарата. Эталонная траектория летательного аппарата задана движением псевдообъекта, называемого далее «цель», и представлена последовательностью из n опорных точек (x_i, y_i) , $i = 1, \dots, n + 1$, $x_1 = x_{n+1}, y_1 = y_{n+1}$. Начальная и конечная точки совпадают. Индексом p будем помечать переменные БПЛА, а индексом e – цели. В простейшем случае цель осуществляет прямолинейное движение между соседними опорными точками траектории с постоянной (эталонной) скоростью $v^{(e)}$ и углом $\theta_i^{(e)}$, задающим направление. БПЛА преследует ее, руководствуясь выбранной стратегией и возможностью управления скоростью и углом тангажа. Известно желаемое время прохождения отдельных опорных точек $t_i^{(e)}$, $i = 1, \dots, n$ и всего маршрута в целом $T^{(e)}$. В результате ветровой нагрузки возможно отклонение летательного аппарата от своего маршрута, причем существенное, что может отклонение реального времени полета $T^{(p)}$ от заданного $T^{(e)}$. Направление ветра может быть как встречным, так и попутным в зависимости от фазы полета.

Скорость БПЛА устанавливается соответствующим рычагом управления в зависимости от ситуации. В дальнейшем ограничимся наличием у цели одной скорости $v^{(e)}$, а у БПЛА двух скоростей $v_1^{(p)} = v^{(e)}$ и $v_2^{(p)} > v^{(e)}$. Угол тангажа $\theta^{(p)}$ устанавливается рычагами управления БПЛА в допустимых пределах $[\theta_{\min}^{(p)}, \theta_{\max}^{(p)}]$. Во время преследования догоняющий «видит» цель, т.е. знает координаты свои и цели в каждый текущий момент времени.

Приведем общую постановку задачи следования БПЛА по заданному маршруту как задачи преследования цели. Пусть $X(t)$ и $Y(t)$ – координаты БПЛА и цели, а $d(X(t), Y(t))$ – расстояние между ними в момент времени t . Задача заключается в построении такого управления $U(t) = (v^{(p)}(t), \theta^{(p)}(t))$ на временном отрезке $[0, T^{(p)}]$ при наличии возмущений и ограничений на управление,

$$\text{что } \int_{t=0}^{T^{(p)}} d(X(t), Y(t)) dt \rightarrow \min.$$

Предлагается решение этой задачи, основанное на действиях, имитирующих поведение летчика (человека-оператора), и заключающееся в выборе стратегий, реализуемых наборами правил в условиях установленных ограничений на управление и действующих возмущений.

2. Решение задачи коммивояжера на основе ИНС

Применительно к задаче коммивояжера нейронная сеть Кохонена использует принцип отображения двумерного распределения n опорных точек (нейронов-кластеров) на одномерный кольцевой маршрут, предварительно заполненный равномерно распределенными вспомогательными точками. После завершения процесса обучения нейронной сети положение кластера в маршруте определится положением его образа в кольцевом выходном слое. Ее использование, как и в случае использования других лучших приближенных методов оптимизации, требует вычислительных затрат, оцениваемых как n^2 . Задача решается в следующем порядке. Сначала в сети фиксируется начальное размещение всех n опорных точек-кластеров с нормированными географическими координатами $(x_i, y_i), i = 1, \dots, n$. Эти значения интерпретируются в сети Кохонена как весовые коэффициенты настройки $(w_{1i}(t), w_{2i}(t)), i = 1, \dots, n$ и определяют размещение точек-кластеров на плоскости в момент времени t . Затем на вход сети в случайном порядке подается m точек с координатами, полученными на основе первоначального равномерного размещения точек на окружности, причем $m \gg n$. Эти координаты интерпретируются как входной вектор $(x_j^{(s)}, y_j^{(s)}), j = 1, \dots, m$, где индекс s подчеркивает расположение точек на окружности. Все точки окружности для удобства работы должны быть организованы в виде списка. Центр и радиус круга подбираются экспериментально. В простейшем случае – это минимальный радиус описанной вокруг точек окружности. Для решения задачи используется типовой алгоритм [Хачумов, 2013] настройки сетевой модели, который приводит к «вытягиванию» точек на окружность. Могут быть также использованы некоторые модификации настройки сети Кохонена, заключающиеся во введении функции окрестности для случайно выбранной точки круга, с тем, чтобы одновременно корректировать положение n точек-кластеров, в зависимости от их расстояния от искомой точки.

Цикл обучения продолжается для всех элементов входного множества точек круга до достижения системой целевого состояния. Пример нескольких последовательных этапов работы ИНС Кохонена показан на рис. 1.

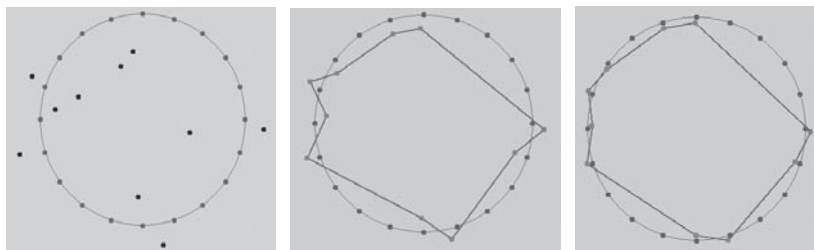


Рис. 1. Начальный, промежуточный и конечный этапы работы ИНС Кохонена

Одной из подзадач при решении задачи коммивояжера с использованием самоорганизующейся карты Кохонена является извлечение маршрута облета после того, как все точки-кластеры лягут на окружность. В идеальном случае после работы ИНС каждая точка-кластер располагается точно на окружности. Но чаще возникает ситуация, что веса не равны координатам. Тогда применяется следующий алгоритм:

Алгоритм. Процедура восстановления маршрута

1. Для всех точек окружности выполнить:
 - найти ближайшую точку-кластер
 - если* кластер не привязан к какой-либо точке окружности, то привязать
 - иначе* удалить точку окружности
2. Для всех точек-кластеров выполнить:
 - если* кластер не привязан к какой-либо точке окружности, создать новую точку и привязать к ней точку-кластер
 - найти ближайшую к кластеру точку на окружности
 - вставить новую точку до или после ближайшей, в зависимости от того, какой путь короче

После применения данного алгоритма можно получить некоторый путь, который принимается за основу построения траектории полета.

3. Управление, основанное на правилах

Схема системы управления, основанной на производственных правилах, была рассмотрена автором в работе [Абрамов и др., 2013]. Существенную роль здесь играет выбор стратегии управления на основе текущих данных о состояниях БПЛА и цели. Рассмотрим некоторые стратегии управления БПЛА.

Стратегия 1 (стратегия движения по точкам) для БПЛА заключается в коррекции движения по текущему отклонению от эталонной траектории следования цели. При этом должны быть пройдены все точки с минимальным отклонением по времени $\Delta^{(T)} = |T^{(p)} - T^{(e)}| \rightarrow \min$.

Стратегия 2 (стратегия сближения с целью) заключается в преследовании эталонной цели. Стратегия осуществляет параллельное сближение [Петросян, 1997] БПЛА с эталонной целью и предполагает вычисление угла тангажа для прогнозирования места встречи с целью [Абрамов и др., 2013]. Здесь точного прохождения летательного аппарата через опорные точки не требуется.

Приведем применяемые производственные правила:

Правило 1 (режим следования): если отклонение и время отработки отклонения для БПЛА и цели не превышают наперед заданных пороговых значений, то применить стратегию 1.

Правило 2 (режим упреждения): если отклонение или время отработки отклонения для БПЛА и цели превышают пороговые значения, то применить стратегию 2.

Значения порогов для смены режима устанавливаются заранее с учетом ветровых нагрузок в данной местности.

4. Моделирование движения БПЛА по заданному маршруту в неспокойной среде

Проведем моделирование полета БПЛА по маршруту, построенному нейронной сетью Кохонена, в условиях ветровой нагрузки с помощью программного обеспечения MATLAB и системы Simulink.

Примем скорость цели $v^{(e)}=30$ м/с, воздушную скорость БПЛА в режиме следования $v_1^{(p)}=30$ м/с, в режиме упреждения (параллельного сближения) $v_2^{(p)}=45$ м/с. Ветровая нагрузка состоит в общем случае из двух составляющих. Первая учитывает распре-

деление скорости ветра по высоте, и ее закон устанавливается в зависимости от местности. Вторая составляющая носит случайный характер и непосредственно влияет на изменение угла тангажа. В зависимости от фазы полета направление ветра может быть как встречным, так и попутным. Среднеквадратичные значения продольной составляющей ветра $\sigma_{\tau} = 3$ м/с. На рис. 2а представлен результат работы ИНС по формированию маршрута облета 20 опорных точек. Видно, что на небольшом участке маршрута имеет место проблема с выбором порядка облета близких по расположению точек, которая может быть решена локальным перебором. На рис. 2б представлены траектории движения эталонной цели (светлая кривая) и БПЛА (темная кривая). Пройденное расстояние измеряется в километрах.

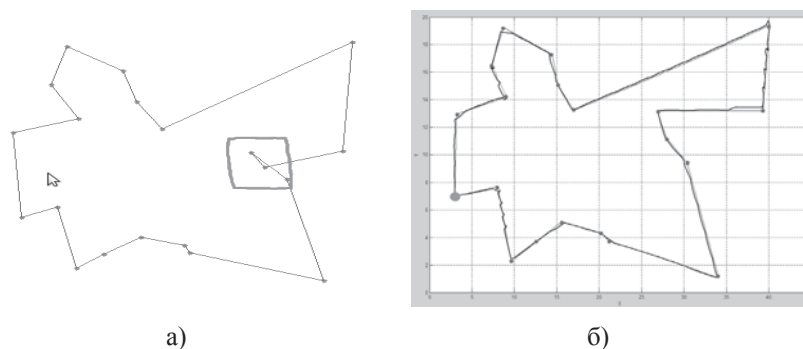


Рис. 2. Результаты моделирования задачи облета территории

Из рис. 2б видно, что БПЛА осуществляет логичные («разумные») действия, имитируя действия пилота по преследованию цели, не сильно отклоняется от заданной траектории и завершает полет в первоначальной точке (выделена утолщением). Общее время моделирования составило 400 секунд.

Заключение

В настоящей работе планирование траектории движения БПЛА рассматривается как решение задачи коммивояжера. Использование ИНС Кохонена позволяет получать приемлемое решение выбора маршрута для задач большой размерности в условиях дефицита

времени и необходимости оперативных действий. В основу решения задачи следования БПЛА по заданному маршруту в возмущенной среде в работе положено применение трех принципов:

– принцип преследования цели, совершающей идеальное движение по маршруту в соответствии с полетным заданием, в условиях воздействий ветровых нагрузок на БПЛА;

– принцип приближенного определения точки встречи БПЛА с целью, необходимого для вычисления угла тангажа и реализации стратегии параллельного сближения;

– принцип имитации действия человека-оператора, управляющего БПЛА, на основе выбора стратегии, реализуемой простыми продукционными правилами.

Проведено моделирование полета БПЛА по заданному маршруту при наличии ветровой нагрузки в соответствии с выбранной стратегией, направленной на минимизацию рассогласования БПЛА и идеальной цели, которое показало, что система управления, основанная на продукционных правилах, достаточно уверенно отрабатывает полетное задание. Система управления быстро реагирует на отклонения от цели и вырабатывает команды по сближению с идеальной траекторией.

СПИСОК ЛИТЕРАТУРЫ

[Хачумов, 2010] Хачумов М.В. Планирование траектории движения летательного аппарата // Искусственный интеллект и принятие решений, №4, 2010, с. 70–76.

[Кудрявцев, 1984] Кудрявцев Е.М. Исследование операций в задачах, алгоритмах и программах. – М.: Радио и связь, 1984.

[Абрамов и др., 2013] Абрамов Н.С., Хачумов М.В. Моделирование проводки по маршруту беспилотного летательного аппарата как задачи преследования цели // Авиакосмическое приборостроение, № 9, 2013, с. 9–22.

[Кузин, 1973] Кузин Л.Т. Основы кибернетики. Т. 1. Математические основы кибернетики – М.: Энергия, 1973. – 504 с.

[Ежов и др., 1998] Ежов А.А., Шумский С.А. Нейрокомпьютеринг и его применения в экономике и бизнесе. Нейросетевая оптимизация. – <http://www.intuit.ru/studies/courses/2255/139/lecture/20567>.

[Хачумов, 2013] Хачумов М.В. Сетевая модель кластерного анализа // Прикладная физика и математика, №5, 2013, с. 85–91.

[Петросян, 1997] Петросян Л.А. Дифференциальные игры преследования. – Л.: Изд-во Ленингр. ун-та. 1997.

УДК 004.4, 004.89, 004.932

**РАСПРЕДЕЛЕННАЯ СИСТЕМА УПРАВЛЕНИЯ
КОАЛИЦИЯМИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ
АППАРАТОВ, ФУНКЦИОНИРУЮЩИХ В УСЛОВИЯХ
ОТСУТСТВИЯ СИГНАЛОВ ГЛОБАЛЬНЫХ
НАВИГАЦИОННЫХ СИСТЕМ¹**

К.С. Яковлев (*yakovlev@isa.ru*)

Институт системного анализа Российской академии наук, Москва

В.В. Хитков (*vskhitkov@gmail.com*)

М.И. Логинов (*lunarstrainut@gmail.com*)

А.В. Петров (*gmdidro@gmail.com*)

Рыбинский государственный авиационный технический университет
имени П.А. Соловьева, Рыбинск

В работе описывается распределенная система управления группой беспилотных летательных аппаратов в среде без доступа к GPS/ГЛОНАСС на примере коллективной навигации, локализации маркеров и построения карты маркеров в помещении группой квадрокоптеров AR.Drone 2.0. Разработанная интеллектуальная система управления основана на модели 4D/RCS и реализована с применением платформы ROS. Задача картирования решается с помощью реализации алгоритма EKF-SLAM.

Введение

Современные интеллектуальные системы характеризуются повышением степени автономности и распределенности. С одной стороны, автономность повышается с точки зрения независимости от оператора, а с другой стороны, в случае многоагентных систем, речь идет об автономности как степени автоматизации агентов, выполняющих единую задачу. Распределенность современных систем понимается и как обеспечение коллективного взаимодействия между несколькими интеллектуальными агентами (распределение

¹ Работа выполнена при финансовой поддержке ОНИТ РАН, проект «Развитие методов интеллектуального анализа данных и управления робототехническими системами с применением бесконтактных человеко-машинных интерфейсов».

знаний и информации между агентами), и с точки зрения технологий распределенного программного обеспечения.

В статье представлена распределенная система управления и навигации беспилотных летательных аппаратов (БЛА), в частности, квадрокоптеров, функционирующих в условиях отсутствия сигналов глобальных навигационных систем (GPS/ГЛОНАСС). Реализация представленной системы демонстрируется на примере задачи построения карты, решаемой группой БЛА. N квадрокоптеров (использовались AR.Drone 2.0) функционируют в закрытом помещении и управляются удаленно (по Wi-Fi) программной системой управления с наземной станции (ноутбука с ОС Linux). В помещении расположены несколько маркеров – кубиков с нанесенными на стороны QR-кодами. Каждый БЛА ищет маркеры и вычисляет их положения относительно собственной стартовой позиции (строит локальную карту), при этом параллельно этому локальные карты от различных БЛА интегрируются в глобальную и уточняются с течением времени.

В статье предлагается концептуальная структура архитектуры системы управления, описываются реализованные алгоритмы распознавания маркеров, вычисления координат маркеров и построения карты. Также рассматривается реализация системы управления на базе платформы ROS (Robotic Operation System). В заключение приводятся результаты эксперимента и дальнейшие направления работы.

1. Многоуровневая архитектура распределенной системы управления беспилотными летательными аппаратами

К настоящему моменту известно множество подходов к построению систем управления сложными техническими объектами, в том числе беспилотной техникой и, в частности, беспилотными летательными аппаратами [Осипов и др., 2009], [Яковлев и др., 2013]. Одним из наиболее распространенных является подход, основанный на функциональной декомпозиции, когда система управления представляет собой совокупность модулей, каждый из которых предназначен для решения определенного типа задач. Обычно выделяют следующие задачи и соответствующие им модули системы управления: построение модели (картины) мира, концептуальное планирование (планирование миссии), планирование траектории и

навигация, обеспечение взаимодействия, парирование нештатных ситуаций, управление коммуникацией, управление датчиками и актуаторами БЛА [Яковлев и др., 2013]. Пример системы управления беспилотными летательными аппаратами, содержащей модули для решения вышеперечисленных задач и состоящей из одного уровня, приведен, например, в [Jameson et al., 2005].

Стоит заметить, что наиболее распространено разбиение функциональности системы управления не только по отдельным модулям, но и по отдельным уровням. Обычно, когда речь идет о создании интеллектуальных беспилотных транспортных средств, выделяют три уровня управления: стратегический (верхний), тактический (средний) и реактивный (нижний). К примерам таких архитектур можно отнести ЗТ [Bonasso et al., 1995], ATLANTIS [Gat, 1992] и др. Основным отличием модулей, расположенных на разных уровнях, являются степень обработки входящей информации и время отклика [Яковлев и др., 2013]. Фактически уровень в современных системах управления беспилотной техникой есть абстракция «интеллектуальности»: чем выше уровень, тем более «интеллектуальные» задачи на нем решаются.

Система управления для решения поставленной задачи имеет ряд особенностей, оказывающих непосредственное влияние на ее архитектуру. Во-первых, отсутствует необходимость реализовывать стратегический уровень управления, т.к. рассматривается только одна высокоуровневая задача – «построение карты». Во-вторых, для решения задачи использовались БПЛА Parrot AR.Drone, каждый из которых оснащен бортовым вычислителем с предустановленным программным обеспечением, *уже* реализующим функционал модулей управления реактивного уровня. Методы и алгоритмы реактивного управления реализованы компанией Parrot и описаны в [Bristeau et al., 2011]. Таким образом, отсутствует необходимость реализации реактивного управления: имеющееся ПО полностью автоматизирует управление актуаторами (двигателями) БЛА для обеспечения выполнения следующих элементарных маневров: взлет, посадка, поворот по/против часовой стрелки, полет вперед/назад и вправо/влево. Команды на осуществление этих маневров передаются на бортовой вычислитель по wi-fi каналу от управляющей станции (ноутбука). Таким образом, основное внимание было уделено разработке модулей тактического уровня управления, от-

ветственных за решения таких задач, как распознавание маркеров по видеопотоку, построение локальных карт, интеграции локальных карт в глобальную карту (модель) местности и ряд других.

При реализации тактического уровня управления за основу были взяты формализм и модель архитектуры системы управления беспилотными транспортными средствами и их коалициями 4D/RCS [Albus et al., 2002], разработанные группой профессора Албуса в Национальном институте стандартов США (NIST). Особенность этой модели – абстрактно-функциональная декомпозиция, а именно: согласно 4D/RCS каждый уровень СУ состоит из совокупности узлов, каждый из которых в свою очередь есть совокупность четырех процессов:

- обработка сенсорной информации (ОСИ);
- построение модели мира (ПММ);
- оценивание свойств (ОС);
- выбор управляющих воздействий (генерация поведения) (ГП).

Состав компонент узлов 4D/RCS не меняется от уровня к уровню, но меняется семантика протекающих в узлах процессов – например, на более высоком уровне управления под генерацией поведения понимается планирование поведения (действий), а на более низком – планирование траектории (перемещений).

В рассматриваемом случае 4D/RCS узел – абстракция отдельного объекта управления (квадрокоптера Parrot AR.Drone), а соответствующие процессы имеют следующую интерпретацию:

ОСИ – получение (по wi-fi каналу) и транскодирование видеопотоков от двух видеокамер, установленных на БЛА; получение (также по wi-fi) данных, касающихся состояния и ориентации БЛА (AT-пакетов, содержащих результаты измерений датчиков БЛА); интеграция данных.

ОС – определение ориентации БЛА в глобальной системе координат; распознавание и идентификация маркеров; определение расстояния до маркеров (как в локальной, так и глобальной системе координат).

ПММ – построение локальной карты, интеграция множества локальных карт (карт, построенных различными БЛА) в единую, глобальную модель (карту) местности.

ГП – выбор очередного промежуточного пункта маршрута полета с помощью системы правил.

Как было сказано выше, каждый узел 4D/RCS отвечает за управление одним БЛА. Функциональные процессы на программном уровне реализованы как модули Robotic operating system [Quigley et al., 2009] (подробное описание см. ниже), исполняемые параллельно в операционной системе Linux. Система управления обеспечивает механизмы поддержки глобальной базы данных, т.е. локальные карты, построенные отдельными БЛА, хранятся на управляющей станции и обрабатываются с целью их дальнейшей интеграции в единую глобальную карту местности (доступную всем БЛА).

Ниже будут более подробно описаны подходы, методы и алгоритмы, используемые для определения ориентации БЛА в пространстве, идентификации и распознавания маркеров, картирования. Также будут описаны существенные детали программной реализации системы управления.

2. Алгоритм построения карты расположения маркеров

В ходе работы реализован алгоритм построения карты расположения маркеров, который может применяться в практических задачах автономной навигации при возможности установки визуальных маркеров. Карта строится одновременно несколькими БЛА, каждый из которых наблюдает некоторое подмножество маркеров, расположенных в общей среде функционирования, с последующим их объединением в общее множество (глобальную карту маркеров). Это происходит следующим образом.

Каждый БЛА (имеется в виду узел программы управления конкретным БЛА на наземной станции – см. раздел 3) во время полета использует видеопоток и информацию инерциальных датчиков для одновременного вычисления своего положения и построения карты с помощью EKF-SLAM-алгоритма одновременной навигации и построения карты с использованием расширенного фильтра Калмана). Каждый БЛА независимо локализует видимые маркеры на кадрах видеопотока, идентифицирует их (они полагаются уникальными) и вычисляет их относительные координаты по 6 степеням свободы (перемещение и поворот) с помощью компьютерного зрения. Для каждого найденного маркера проверяется его наличие на глобальной карте, копия которой доступна узлу управления БЛА в любой момент. При этом возможны следующие ситуации.

1) Маркер не присутствует на глобальной карте, а БЛА не знает своего расположения в глобальной системе координат (т.е. он еще не наблюдал тех же маркеров, что и другие БЛА). В этом случае он наносится на индивидуальную карту БЛА с координатами, усредненными за несколько наблюдений в разных кадрах.

2) Маркер присутствует на глобальной карте, но БЛА еще не знает своих координат в глобальной системе. В этом случае маркер используется как связующее звено между локальной системой координат данного БЛА и глобальной системой координат. Осуществляется единовременная трансляция локальной системы координат БЛА к глобальной с пересчетом координат всех объектов на индивидуальной карте БЛА, которая в дальнейшем совпадает с глобальной.

3) Маркер не присутствует на глобальной карте, а БЛА знает свое расположение в глобальной системе координат. В этом случае его карта совпадает с глобальной картой, и маркер вместе с его координатами наносится, таким образом, и на индивидуальную, и на общую карту. Координаты фильтруются за несколько наблюдений.

4) Маркер присутствует на глобальной карте, и БЛА знает свое местоположение. Этот случай используется для уточнения координат как БЛА, так и устранения «перекоса» карты, с помощью «блочного уравнивания» (bundle adjustment) по набору ключевых положений, подобно [Klein et al., 2007] (этот подход применим ввиду полагаемого небольшого количества маркеров-кубиков).

В качестве системы маркеров и инструмента их локализации, идентификации и определения положения по 6 степеням свободы в работе используется библиотека ArUco [Munoz-Salinas et al., 2014], предназначенная для распознавания двухмерных штрих-кодов и характеризующаяся надежностью распознавания за счет применения кода Хэмминга. При известных внутренних калибровочных параметрах камеры становится возможным построение карты в метрической системе координат.

Для оценки трехмерного положения БЛА применяется расширенный фильтр Калмана с вектором состояния $(x_b, y_b, z_b, \alpha_b, \beta_b, \gamma_b)^T$, элементы которого соответственно означают 3 пространственные координаты и 3 угла Эйлера в глобальной или индивидуальной системе координат. Предсказание положения на основе показаний инерциальных датчиков выполняется аналогично описанному в

[Engel et al., 2012]. Для оценки координат по видеопотоку используются рассчитанные абсолютные координаты всех локализованных на кадре и уже присутствующих на карте маркеров в конкретный момент времени.

3. Базирующаяся на ROS программная реализация системы управления

Программная реализация описываемой системы основана на платформе ROS и архитектуре 4D/RCS. На рисунке 1 показана структура программы.

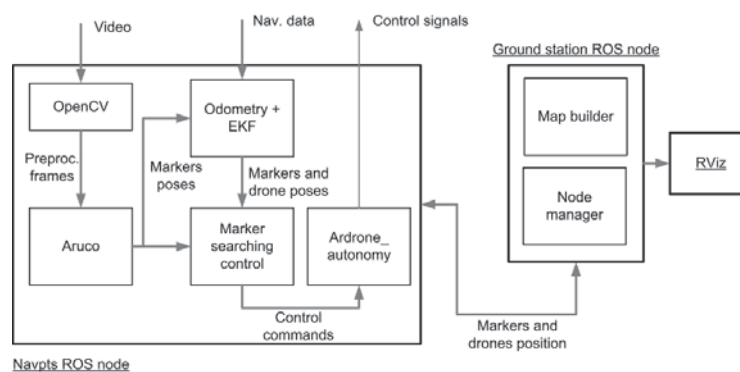


Рис. 1. Структура программной реализации системы распределенного управления, базирующейся на ROS

Пакеты, узлы ROS, а также внешние библиотеки показаны на рисунке 1. Существует два основных узла ROS: узел «Navpts» и «Ground station». Некоторое количество узлов «Navpts» исполняется параллельно, при этом обеспечивая контроль за одним AR.Drone каждый. Узел «Ground Station», запущенный в единственном экземпляре, реализует построение карты и алгоритм ее совместного использования. Пользуясь архитектурой 4D/RCS при описании узла «Navpts», можно отметить, что библиотека OpenCV выполняет обработку сенсорной информации, пакеты Aruco и EKF – оценку значений, пакет поиска маркера вырабатывает требуемое поведение. Процесс построения окружающего мира 4D/RCS распределен между представлениями карты узлами «Navpts» и «Ground Station».

Узел «Ground Station» представляет собой модуль построения карты, реализуемый описанные алгоритмы, и менеджер узлов, запускающий и связывающий необходимые темы (топики) для узлов «Navpts».

Для визуализации карты используется пакет ROS - RViz.

Было выполнено тестирование разработанной реализации системы управления в лабораторных условиях. На рисунке 2, справа, показан результат в форме совместной карты.

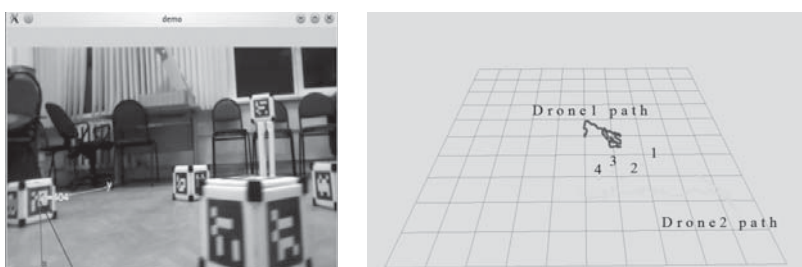


Рис. 2. Обнаружение маркера на изображении (слева) и совместная карта в RViz (справа)

На рисунке 2 (слева) показаны экспериментальная установка и обнаружение маркера. На рисунке 2 (справа) показаны совместная карта с траекториями БПЛА и позициями общих маркеров.

Заключение

В заключение отметим, что предложенная в работе реализация алгоритма построения карты не является оптимальным решением для задачи подобного рода по сравнению с аналогами и используется в качестве иллюстрации выбранной архитектуры и программной реализации системы распределенного управления и навигации БЛА в среде без доступа к данным глобальных систем позиционирования (GPS, ГЛОНАСС). Планируется расширить имеющуюся систему в следующем направлении: интеграция с более эффективным SLAM алгоритмом, который может быть применен в среде с неизвестными ландшафтными метками; модификация процесса генерации поведения путем введения совместного планирования для того, чтобы обеспечить решение в ситуации, в которой один из

БЛА не в состоянии обнаружить ни один маркер; повышение вычислительной эффективности системы.

СПИСОК ЛИТЕРАТУРЫ

[**Осипов и др., 2009**] Осипов Г.С., Тихомиров И.А., Хачумов В.М., Яковлев К.С. Интеллектуальное управление транспортными средствами: стандарты, проекты, реализации // *Авиакосмическое приборостроение*, № 6, 2009. С. 34–43.

[**Jameson et al., 2005**] Jameson S., Franke J., Szczerba R., Stockdale S. Collaborative Autonomy for Manned/Unmanned Teams. AHS International Forum 61. Grapevine. TX. 2005.

[**Яковлев и др., 2013**] Яковлев К.С., Макаров Д.А., Панов А.И., Зубарев Д.В. Принципы построения многоуровневых архитектур систем управления беспилотными летательными аппаратами // *Авиакосмическое приборостроение*, 4, 2013. С.10–28.

[**Bonasso et al., 1995**] Bonasso, R.P., et al. Experiences with an Architecture for Intelligent, Reactive Agents. International Joint Conference on Artificial Intelligence. 1995.

[**Gat, 1992**] Gat, E. Integrating planning and reacting in a heterogenous asynchronous architecture for controlling real-world mobile robots. National Conference for Artificial Intelligence. 1992.

[**Bristeau et al., 2011**] Bristeau, P. J., Callou, F., Vissière, D., & Petit, N. The navigation and control technology inside the AR.Drone micro UAV. In 18th IFAC World Congress (Vol. 18, No. 1, pp. 1477–1484). 2011.

[**Albus et al., 2002**] Albus, J., Huang, H. M., Messina, E., Murphy, K., Juberts, M., Lacaze, A., & Finkelstein, R. 4D/RCS Version 2.0: A reference model architecture for unmanned vehicle systems. National Institute of Standards and Technology, Gaithersburg, MD, NISTIR, 6912. 2002.

[**Quigley et al., 2009**] Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., & Ng, A. Y. ROS: an open-source Robot Operating System. In ICRA workshop on open source software (Vol. 3, No. 3.2). 2009.

[**Klein et al., 2007**] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In International Symposium on Mixed and Augmented Reality, 2007.

[**Munoz-Salinas et al., 2014**] R. Munoz-Salinas, S. Garrido-Jurado, Aruco library, <http://sourceforge.net/projects/aruco/>.

[**Engel et al., 2012**] J. Engel, J. Sturm, D. Cremers. Camera-Based Navigation of a Low-Cost Quadrocopter. In Proc. of the International Conference on Intelligent Robot Systems (IROS), 2012.

УДК 623.746

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ДЛИН ЗВЕНЬЕВ КРЫЛА ОРНИТОПТЕРА НА ЕГО КИНЕМАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ

С.Ф. Яцун (*teormeh@inbox.ru*)

Юго-Западный государственный университет, Курск

С.В. Ефимов (*sclione@yandex.ru*)

Воронежский институт ГПС МЧС России

Г.С. Наумов (*gregory-XXI@yandex.ru*)

Юго-Западный государственный университет, Курск

В данной статье рассматривается орнитоптер – бионический робот с машущими крыльями, имитирующий движения птицы во время полета. Построена математическая модель объекта, представлены результаты в виде графиков. Исследовано влияние длин звеньев крыльев на среднюю скорость робота.

Введение

Имитация или копирование принципов поведения живых организмов позволяет создавать устройства, обладающие уникальными свойствами, нацеленными на повышение уровней комфорта, безопасности и автоматизации жизни современных людей. К устройствам подобного плана относятся орнитоптеры – летающие роботы, имитирующие движения птиц в полете [1]. В данной работе исследуется влияние длин звеньев крыла орнитоптера на его кинематические характеристики.

1. Описание орнитоптера

Рассматриваемый орнитоптер (рис. 1) состоит из корпуса 1 и крыльев, состоящих, в свою очередь, из двух звеньев – 2,3 и 4,5. Для дальнейшего изучения движения такого объекта в полете воспользуемся связанной (относительной) системой координат $O'x'y'z'$, начало которой совпадает с центром тяжести корпуса орнитоптера C_1 . Ось $O'x'$ такой системы координат направлена параллельно продольной оси корпуса орнитоптера, ось $O'y'$ направлена перпендикулярно плоскости крыльев и оси $O'x'$. Плоскость $O'x'y'$

является плоскостью симметрии орнитоптера и делит его на две равные части [2]. Ось $O'z'$ направлена перпендикулярно плоскости симметрии орнитоптера. При всех изменениях положения орнитоптера относительно земной (абсолютной) системы координат $Oxyz$, как линейных, так и угловых, связанная система координат перемещается и вращается вместе с ним.

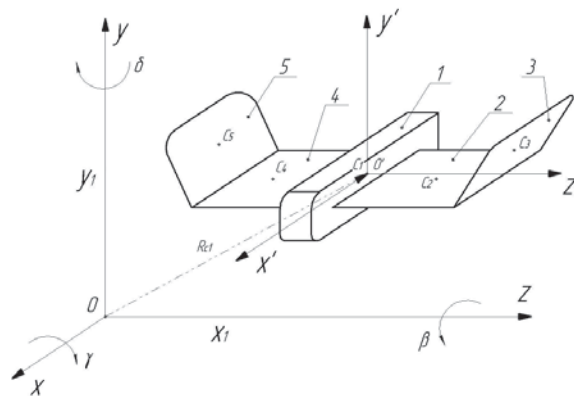


Рис. 1. Расположение орнитоптера в пространстве

В целях упрощения далее будем рассматривать орнитоптер в плоскости Oyz абсолютной системы координат. В точках соединения звеньев установлены приводы, позволяющие звеньям свободно вращаться относительно друг друга. Положение подобной модели орнитоптера на плоскости полностью определяется координатами z_1 и y_1 , а конфигурация звеньев – углами $\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5$ [3–4].

2. Математическая модель орнитоптера

Покажем силы, действующие на орнитоптер во время полета (рис. 2). Под R_1, R_2, R_3, R_4, R_5 указаны полные аэродинамические силы, действующие на звено, а под $R'_1, R'_2, R'_3, R'_4, R'_5$ – боковые составляющие аэродинамических сил, возникающие за счет обтекания крыла воздушным потоком. В точках соединения звеньев установлены приводы, создающие крутящие моменты $M_{21}, M_{32}, M_{41}, M_{54}$, позволяющие звеньям двигаться относительно друг друга.

В целях упрощения процедуры моделирования движения объекта примем следующие допущения:

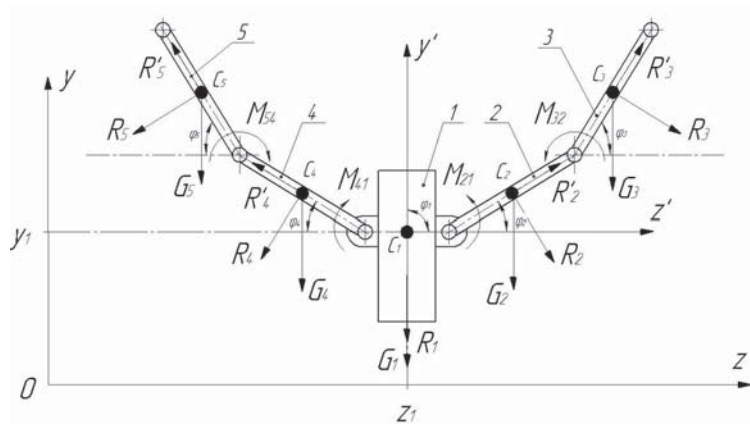


Рис. 2. Математическая модель орнитоопера с учетом действующих сил

1) все звенья летательного аппарата – абсолютно твердые недеформируемые тела;

2) каждое из звеньев представляет собой стержень длиной l_i и массой m_i , сосредоточенной в центре симметрии звена C_i ;

3) величины боковых составляющих аэродинамических сил $R'_1, R'_2, R'_3, R'_4, R'_5$ настолько малы, что ими можно пренебречь.

Приведенные полные аэродинамические силы R_1, R_2, R_3, R_4, R_5 сосредоточены в центрах тяжести эпюр, распределенных по крылу аэродинамических сил, и рассчитываются как [2]:

$$R_i = C_r \cdot \frac{\rho \cdot V_i^2}{2} \cdot S_i \cdot \text{sign}|V_i|$$

где ρ – массовая плотность воздуха;

V_i – скорость точки приложения приведенной силы R_i относительно воздуха, зависящая от обобщенных координат;

S_i – эффективная площадь звена крыла, зависящая от угловых перемещений;

C_r – безразмерный коэффициент полной аэродинамической силы.

Движение орнитоопера описывается системой из семи дифференциальных уравнений:

$$\begin{aligned}
& \ddot{\phi}_1 (J_1 + \frac{1}{4} m_2 l_1^2 + \frac{1}{4} m_3 l_1^2) + \ddot{\phi}_2 (\frac{1}{2} m_2 + m_3) \frac{l_1 l_2}{2} \cos(\phi_1 - \phi_2) + \ddot{\phi}_3 m_3 \frac{l_1 l_3}{4} \cos(\phi_1 - \phi_3) + \\
& + \ddot{y}_{C1} (m_2 + m_3) \frac{l_1}{2} \cos \phi_1 - \ddot{z}_{C1} (m_2 + m_3) \frac{l_1}{2} \sin \phi_1 + \dot{\phi}_2^2 (\frac{1}{2} m_2 + m_3) \frac{l_1 l_2}{2} \sin(\phi_1 - \phi_2) + \\
& + \dot{\phi}_3^2 m_3 \frac{l_1 l_3}{4} \sin(\phi_1 - \phi_3) = -R_3 \cdot \frac{l_3}{2} - R_2 \cdot \frac{l_2}{2} - R_1 (\cos(\phi_1) + \sin(\phi_1)) - \\
& - m_1 g - m_2 g \frac{l_2}{3} \cos \phi_2 - m_3 g \frac{l_3}{2} \cos \phi_3
\end{aligned}$$

$$\begin{aligned}
& \ddot{\phi}_2 (J_2 + \frac{1}{4} m_2 l_2^2 + m_3 l_2^2) + \ddot{\phi}_1 (\frac{1}{2} m_2 + m_3) \frac{l_1 l_2}{2} \cos(\phi_1 - \phi_2) + \ddot{\phi}_3 m_3 \frac{l_2 l_3}{2} \cos(\phi_2 - \phi_3) + \\
& + \ddot{y}_{C1} (m_2 + 2m_3) \frac{l_2}{2} \cos \phi_2 - \ddot{z}_{C1} (m_2 + 2m_3) \frac{l_2}{2} \sin \phi_2 + \dot{\phi}_1^2 (\frac{1}{2} m_2 + m_3) \frac{l_1 l_2}{2} \sin(\phi_2 - \phi_1) + \\
& + \dot{\phi}_3^2 m_3 \frac{l_2 l_3}{2} \sin(\phi_2 - \phi_3) = M_{21} - R_3 \cdot \frac{l_3}{2} - R_2 \cdot \frac{l_2}{2} - m_2 g \frac{l_2}{2} \cos \phi_2 - m_3 g \frac{l_3}{2} \cos \phi_3
\end{aligned}$$

$$\begin{aligned}
& \ddot{\phi}_3 (J_3 + \frac{1}{4} m_3 l_3^2) + \ddot{\phi}_2 \frac{l_2 l_3}{2} m_3 \cos(\phi_2 - \phi_3) + \ddot{\phi}_1 m_3 \frac{l_1 l_3}{4} \cos(\phi_1 - \phi_3) + \ddot{y}_{C1} m_3 \frac{l_3}{2} \cos \phi_3 - \\
& - \ddot{z}_{C1} m_3 \frac{l_3}{2} \sin \phi_3 + \dot{\phi}_1^2 m_3 \frac{l_1 l_3}{4} \sin(\phi_3 - \phi_1) + \dot{\phi}_2^2 m_3 \frac{l_2 l_3}{2} \sin(\phi_3 - \phi_2) = M_{32} - R_3 \cdot \frac{l_3}{2} - \\
& - m_3 g \cdot \frac{l_3}{2} \cos \phi_3
\end{aligned}$$

$$\begin{aligned}
& \ddot{\phi}_4 (J_4 + \frac{1}{4} m_4 l_4^2 + m_5 l_4^2) + \ddot{\phi}_1 (\frac{1}{2} m_4 + m_5) \frac{l_1 l_4}{2} \cos(\phi_1 - \phi_4) + \ddot{\phi}_3 m_5 \frac{l_4 l_5}{2} \cos(\phi_4 - \phi_5) + \\
& + \ddot{y}_{C1} (m_4 + 2m_5) \frac{l_4}{2} \cos \phi_4 - \ddot{z}_{C1} (m_4 + 2m_5) \frac{l_4}{2} \sin \phi_4 + \dot{\phi}_1^2 (\frac{1}{2} m_4 + m_5) \frac{l_1 l_4}{2} \sin(\phi_4 - \phi_1) + \\
& + \dot{\phi}_3^2 m_5 \frac{l_4 l_5}{2} \sin(\phi_4 - \phi_5) = M_{41} + R_5 \cdot \frac{l_5}{2} + R_4 \cdot \frac{l_4}{2} - m_5 g \frac{l_5}{2} \cos \phi_5 - m_4 g \frac{l_4}{2} \cos \phi_4
\end{aligned}$$

$$\begin{aligned}
& \ddot{\phi}_5 (J_5 + \frac{1}{4} m_5 l_5^2) + \ddot{\phi}_4 \frac{l_4 l_5}{2} m_5 \cos(\phi_4 - \phi_5) + \ddot{\phi}_1 m_5 \frac{l_1 l_5}{4} \cos(\phi_1 - \phi_5) + \ddot{y}_{C1} m_5 \frac{l_5}{2} \cos \phi_5 - \\
& - \ddot{z}_{C1} m_5 \frac{l_5}{2} \sin \phi_5 + \dot{\phi}_1^2 m_5 \frac{l_1 l_5}{4} \sin(\phi_5 - \phi_1) + \dot{\phi}_4^2 m_5 \frac{l_4 l_5}{2} \sin(\phi_5 - \phi_4) = M_{54} - R_5 \cdot \frac{l_5}{2} - \\
& - m_5 g \cdot \frac{l_5}{2} \cos \phi_5
\end{aligned}$$

$$\begin{aligned}
& \ddot{z}_{C1} (m_1 + m_2 + m_3) - \dot{\phi}_1 \frac{l_1}{2} \sin \phi_1 (m_2 + m_3) - \dot{\phi}_2 \frac{l_2}{2} \sin \phi_2 (m_2 + 2m_3) - \dot{\phi}_3 m_3 \frac{l_3}{2} \sin \phi_3 - \\
& - \dot{\phi}_1^2 (m_2 + m_3) \frac{l_1}{2} \cos \phi_1 - \dot{\phi}_2^2 (m_2 + 2m_3) \frac{l_2}{2} \cos \phi_2 - \dot{\phi}_3^2 m_3 \frac{l_3}{2} \cos \phi_3 = -R_1 \cos \phi_1 + \\
& + R_2 \sin \phi_2 + R_3 \sin \phi_3 + R_4 \sin \phi_4 + R_5 \sin \phi_5
\end{aligned}$$

3. Исследование влияния длин звеньев на среднюю скорость движения

Разработанная математическая модель позволила провести исследование влияния длин звеньев крыла орнитоопера на среднюю скорость движения корпуса. Моделирование проводилось для робота массой 1 кг. Массовая плотность воздуха $\rho=1,2 \text{ кг} \cdot \text{с}^2/\text{м}^4$, коэффициент аэродинамической силы $C_R=0,2$.

Характер движения робота определяется многими параметрами – длиной звеньев, амплитудой управляющих моментов, диапазоном углов, в котором осуществляется поворот звеньев. Выберем критерий оценки степени влияния каждого из этих параметров – этим критерием будет средняя скорость V_{C1}^{cp} корпуса робота за два полных цикла движения.

Предварительный анализ показал, что изменение длины первого звена на скорость робота никакого влияния не оказывает, так как моделируемое движение – синхронное, и это звено не имеет углового перемещения. Будем изменять длину второго звена l_2 и оценивать V_{C1}^{cp} .

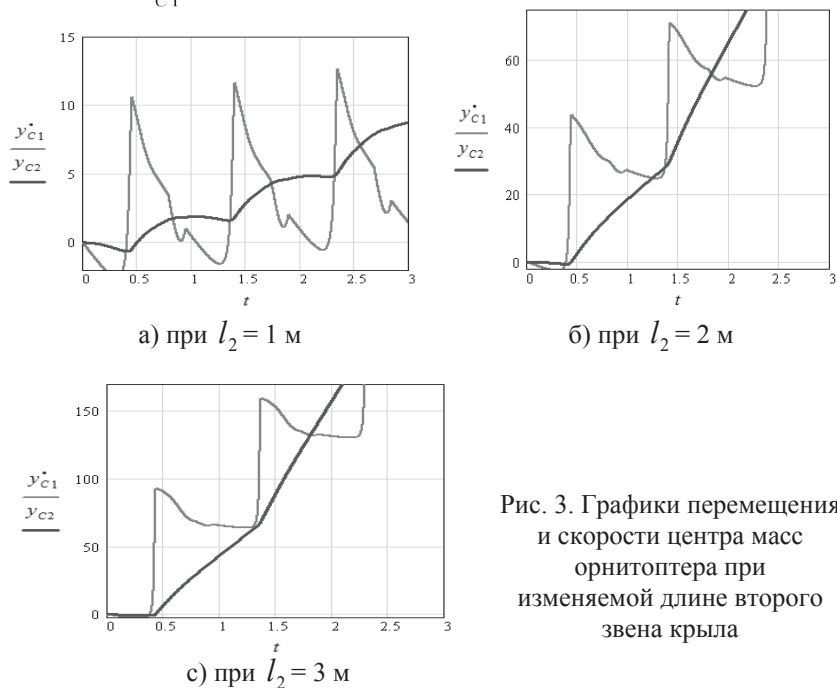


Рис. 3. Графики перемещения и скорости центра масс орнитоопера при изменяемой длине второго звена крыла

Видно, что увеличение длины второго звена в 2 раза ведет к увеличению средней скорости взлета орнитоопера почти в 4 раза, в 3 раза – к увеличению средней скорости почти в 9 раз. Зависимость между этими двумя параметрами прямо пропорциональная, причем явно нелинейная, возможно, носящая квадратичный характер.

Исследуем влияние длины второго звена l_3 на среднюю скорость движения V_{C1}^{cp} .

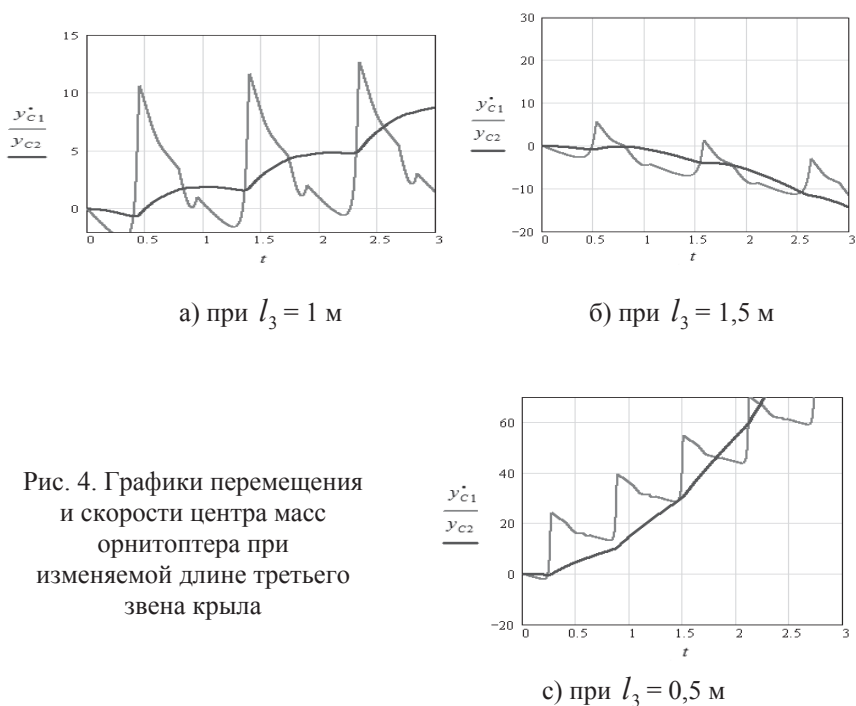


Рис. 4. Графики перемещения и скорости центра масс орнитоопера при изменяемой длине третьего звена крыла

На графиках видно, что наблюдается снижение средней скорости V_{C1}^{cp} при увеличении длины третьего звена, и смена режима «взлет» на режим «посадка» при $l_3 > 1,35$ м. Зависимость между этими двумя параметрами обратно пропорциональная.

Из полученных результатов можно сделать вывод, что для достижения большей эффективности полета длины звеньев 2,3 должны быть разными.

Заключение

В работе рассмотрен орнитооптер, крылья которого состоят из двух звеньев, построена его математическая модель. Была получена система из семи дифференциальных уравнений, описывающая движение многозвенной системы в воздушной среде.

Моделирование движения выполнено для типа полета, при котором корпус робота совершает горизонтальный полет. Получены зависимости соответствующих обобщенных координат от времени, представлены результаты численного моделирования. Исследовано влияние длин звеньев крыла на среднюю скорость движения орнитооптера.

СПИСОК ЛИТЕРАТУРЫ

[Вотяков и др., 1975] Вотяков А.А., Каюнов Н.Т. Аэродинамика и динамика полета самолета – уч. пос. – М.: Издательство ДОСААФ, 1975. – 295 с.

[Нашукевич и др., 1968] Нашукевич А.В., Неволин Ф.А., Немировский Ф.А. Аэродинамика самолета. – М.: «Воениздат», 1968.

[Яцун и др., 2013] Яцун С.Ф., Волкова Л.Ю., Рублев С.Б., Наумов Г.С. Трехзвенный ползающий робот как средство передвижения // Прогресс транспортных средств и систем – 2013: материалы Международной научно-практической конференции / ВолгГТУ – Волгоград, 2013. – Ч. 1. – С. 293–294.

[Локтионова и др., 2013] Локтионова О.Г., Яцун С.Ф., Рублев С.Б., Волкова Л.Ю., Наумов Г.С. Моделирование движения трехзвенного робота с управляемыми силами трения по абсолютно гладкой горизонтальной поверхности // Научный журнал КубГАУ. 2013. № 91(07). С. 41.

УДК 681.51

МОДЕЛИРОВАНИЕ ДВИЖЕНИЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ КВАДРОРОТАЦИОННОГО ТИПА

**С.Ф. Яцун, Н.И. Попов, О.В. Емельянова,
А.И. Савин** (*teormeh@inbox.ru*)

ГОУ ВПО «Юго-Западный государственный университет», Курск,
Россия

В работе рассмотрены вопросы математического моделирования движения квадрокоптера, с учетом массогабаритных свойств четырех электроприводов, снабженных редуктором. Приведена расчетная схема и составлены дифференциальные уравнения на основе общих теорем динамики, которые описывают взаимосвязанные электромагнитные и механические процессы в электромеханической системе приводов винтов квадрокоптера. Предложенные нелинейные дифференциальные уравнения решаются совместно с кинематическими соотношениями, выражающими проекции угловой скорости тела на оси связанной системы координат через угловые скорости углов крена, тангажа и рыскания.

Введение

Квадрокоптер – это летательный аппарат с четырьмя несущими винтами, вращающимися диагонально в противоположных направлениях. Он обладает рядом преимуществ по сравнению с беспилотными аналогами самолетного типа, таких как возможность вертикального взлета и посадки, маневренность в полете, малая взлетная масса при существенной массе полезной нагрузки, надежность и компактность. Благодаря простоте конструкции квадрокоптеры часто используются в любительском моделировании, удобны для недорогой аэрофото- и киносъемки – громоздкая камера вынесена из зоны действия винтов.

Для изучения основных закономерностей движения квадрокоптера, особенностей системы управления, необходимой для стабилизации аппарата в заданном положении и движении центра масс по заданной траектории, необходимо составить математическую модель, описывающую пространственное движение летающего робота.

Математическая модель квадрокоптера

Будем рассматривать квадрокоптер как электромеханическую систему, корпус которой можно моделировать твердым телом с 6-ю степенями свободы [Емельянова и др., 2013a], [Емельянова и др., 2013b], [Попов и др., 2014a], [Попов и др., 2014b], [Bresciani, 2008], [Tahar, 2011].

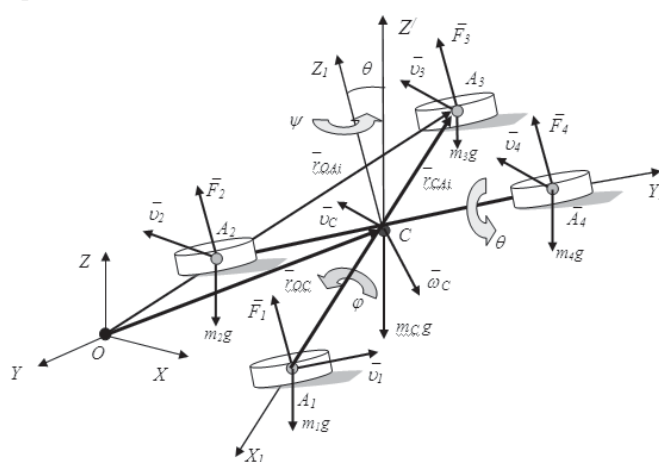


Рис. 1. Расчетная схема квадрокоптера

Пусть положение центра масс квадрокоптера C совпадает с началом подвижной системы координат $CX_1Y_1Z_1$, а в неподвижной декартовой системе координат описывается координатами X, Y, Z (рис. 1).

Ориентацию в пространстве задают углы Эйлера-Крылова, которые обычно применяются в авиационной технике при описании движения аппарата и составляют так называемые углы: крена, тангажа и рысканья [Павловский и др., 1990]. Они соответствуют следующей последовательности поворотов:

1. Поворот на угол ψ относительно вертикальной оси OZ ($R z, \psi$) – рыскание.
2. Поворот на угол θ относительно главной поперечной оси инерции OY ($R y, \theta$) – тангаж.
3. Поворот на угол φ вокруг продольной оси OX ($R x, \varphi$) – крен.

При полете квадрокоптера на него действуют аэродинамические

ские силы несущих винтов $\bar{F}_1, \bar{F}_2, \bar{F}_3, \bar{F}_4$, приложенные к центрам масс роторов A_1, A_2, A_3, A_4 , причем силы \bar{F}_i параллельны оси CZ_i . Силы тяжести корпуса $m_c g$ и роторов винтов $m_i g$ приложены в точках C и A_i соответственно (рис. 1) [Попов и др., 2014a], [Попов и др., 2014b].

Положение центра масс квадрокоптера определяет радиус-вектор $r_{OC} = [X, Y, Z]^T$.

Условимся в дальнейшем векторы в системы координат $OXYZ$ и $CX_i Y_i Z_i$ обозначать символами $^{(0)}$ и $^{(1)}$ соответственно. Тогда, например, проекции векторов сил в $^{(0)}$ системе координат можно выразить через проекции сил в $^{(1)}$ системе координат с помощью выражения:

$$F_i^{(0)} = T_{10} \cdot F_i^{(1)}, \quad (1)$$

где T_{10} – матрица перехода из $^{(1)}$ в $^{(0)}$ систему координат.

Матрица T_{10} получается путем перемножения трех основных матриц вращения и имеет следующий вид [Bresciani, 2008], [Tahar, 2011], [Павловский и др., 1990]:

$$T_{10} = (\psi, \theta, \varphi) = R(z, \psi) \times R(y, \theta) \times R(x, \varphi) = \begin{bmatrix} \cos \psi \cos \theta & \cos \psi \sin \theta \sin \varphi - \cos \varphi \sin \psi & \sin \psi \sin \theta \sin \varphi + \cos \psi \cos \varphi \sin \theta \\ \sin \psi \cos \theta & \cos \psi \cos \varphi + \sin \psi \sin \theta \sin \varphi & \cos \varphi \sin \psi \sin \theta - \cos \psi \sin \varphi \\ -\sin \theta & \cos \theta \sin \varphi & \cos \varphi \cos \theta \end{bmatrix} \quad (2)$$

Запишем очевидное равенство:

$$r_{OA_i}^{(0)} = \bar{r}_{OC}^{(0)} + \bar{r}_{CA_i}^{(0)}, \quad (3)$$

где

$$\bar{r}_{CA_i}^{(0)} = T_{10} \cdot \bar{r}_{CA_i}^{(1)}, \quad (4)$$

векторы $\bar{r}_{CA_i}^{(1)}$ для точек A_i имеют вид:

$$\bar{r}_{CA_1}^{(1)} = \begin{bmatrix} l \\ 0 \\ 0 \end{bmatrix}, \quad \bar{r}_{CA_2}^{(1)} = \begin{bmatrix} 0 \\ -l \\ 0 \end{bmatrix}, \quad \bar{r}_{CA_3}^{(1)} = \begin{bmatrix} -l \\ 0 \\ 0 \end{bmatrix}, \quad \bar{r}_{CA_4}^{(1)} = \begin{bmatrix} 0 \\ l \\ 0 \end{bmatrix}, \quad (5)$$

где l – расстояние от центра масс квадрокоптера C до центра масс роторов A_i .

Скорости точек A_i определим, продифференцировав равенство (3) по времени:

$$\bar{v}_{A_i}^{(0)} = \frac{\bar{r}_{OA_i}^{(0)}}{dt} = \frac{\bar{r}_{OC}^{(0)}}{dt} + \frac{\bar{r}_{CA_i}^{(0)}}{dt}. \quad (6)$$

Производная по времени от равенства (4) дает:

$$\bar{v}_{A_i}^{(0)} = \bar{v}_C^{(0)} + \dot{T}_{10} \cdot \bar{r}_{CA_i}^{(1)}, \quad (7)$$

где $\bar{v}_C^{(0)} = \bar{i}\dot{X} + \bar{j}\dot{Y} + \bar{k}\dot{Z}$ – скорость центра масс квадрокоптера.

Определим количество движения i -ой массы по формуле:

$$\bar{q}_i = m_i \bar{v}_{A_i} = m_i (\bar{v}_C + \dot{T}_{10} \cdot \bar{r}_{CA_i}^{(1)}). \quad (8)$$

Изменение количества движения найдем из выражения:

$$\frac{d\bar{q}_i}{dt} = m_i \left(\frac{d\bar{v}_C}{dt} + \ddot{T}_{10} \cdot \bar{r}_{CA_i}^{(1)} \right) = T_{10} \bar{F}_i^{(1)}. \quad (9)$$

Вектор количества движения рассматриваемой системы, состоящей из корпуса и 4 винтов, определим по формуле:

$$\bar{Q} = m_C \bar{v}_C + \sum_{i=1}^4 m_{A_i} \bar{v}_{A_i}. \quad (10)$$

Запишем теорему об изменении количества движения механической системы в дифференциальной форме:

$$\frac{d\bar{Q}}{dt} = m_C \frac{d\bar{v}_C}{dt} + \sum m_i \left(\frac{d\bar{v}_C}{dt} + \ddot{T}_{10} \cdot \bar{r}_{CA_i}^{(1)} \right) = (m_C + \sum m_i) \frac{d\bar{v}_C}{dt} + \ddot{T}_{10} \sum m_i \bar{r}_{CA_i}^{(1)} = T_{10} \sum \bar{F}_i^{(1)} \quad (11)$$

После соответствующих преобразований получим систему дифференциальных уравнений, описывающих движение центра масс рассматриваемой системы (квадрокоптера).

$$\begin{cases} m\ddot{X} = (\sin \psi \sin \varphi + \cos \psi \cos \varphi \sin \theta) \cdot \sum F_i^{(1)} \\ m\ddot{Y} = (\cos \varphi \sin \psi \sin \theta - \cos \psi \sin) \cdot \sum F_i^{(1)} \\ m\ddot{Z} = \cos \varphi \cos \theta \cdot \sum F_i^{(1)} - mg \end{cases} \quad (12)$$

где $m = m_C + m_i$ – общая масса квадрокоптера.

Рассмотрим угловые скорости вращения роторов квадрокоптера в локальной системе координат $CX_i Y_i Z_i$ (рис. 2). Для этого введем систему координат $A_i x_i y_i z_i$, которая совпадает с центром масс m_i роторов.

Вектор абсолютной угловой скорости вращения i -ого ротора определим по формуле:

$$\begin{aligned} \bar{\Omega}_i &= \bar{\omega}_i + \bar{\omega}_C, \quad i = 1 \dots 4; \\ \bar{\omega}_i &= \bar{i}_i \omega_{ix} + \bar{j}_i \omega_{iy} + \bar{k}_i \omega_{iz}; \end{aligned} \quad (13)$$

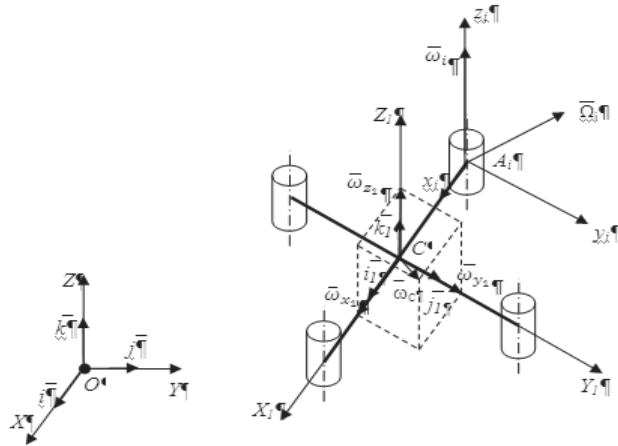


Рис. 2. Расчетная схема определения кинетического момента квадрокоптера

$$\bar{\omega}_C = \bar{i}_1 \omega_{CX_1} + \bar{j}_1 \omega_{CY_1} + \bar{k}_1 \omega_{CZ_1}, \quad (14)$$

где i_1, j_1, k_1 и i_i, j_i, k_i – единичные векторы системы координат $CX_1Y_1Z_1$ и $A_ix_jz_i$, $\bar{\Omega}_i$ – абсолютная угловая скорость вращения i -ого ротора в системе координат $CX_1Y_1Z_1$; $\bar{\omega}_C, \bar{\omega}_i$ – векторы угловых скоростей вращения корпуса и i -ого ротора в системе координат $CX_1Y_1Z_1$ и $A_ix_jz_i$ представим в виде:

$$\bar{\omega}_C = \begin{vmatrix} \omega_{X_1} \\ \omega_{Y_1} \\ \omega_{Z_1} \end{vmatrix}, \quad \bar{\omega}_i = \begin{vmatrix} 0 \\ 0 \\ \omega_i \end{vmatrix}, \quad \bar{\Omega}_i = \begin{vmatrix} \omega_{x_1} \\ \omega_{y_1} \\ \omega_i + \omega_{Z_1} \end{vmatrix} \quad (15)$$

или
$$\bar{\Omega}_i = \bar{i}_1 \Omega_x + \bar{j}_1 \Omega_y + \bar{k}_1 \Omega_z. \quad (16)$$

Определим момент количества движения ротора в системе координат $A_ix_jz_i$

$$\bar{L}_{iA_i} = I_{A_i} \bar{\Omega}_i, \quad (17)$$

где $I_{iA_i} = \begin{vmatrix} J_{A_i}^x & 0 & 0 \\ 0 & J_{A_i}^y & 0 \\ 0 & 0 & J_{A_i}^z \end{vmatrix}$ – тензор инерции ротора.

Тогда кинетический момент равен:

$$\bar{L}_{iA_i} = \begin{vmatrix} J_{A_i}^x & 0 & 0 \\ 0 & J_{A_i}^y & 0 \\ 0 & 0 & J_{A_i}^z \end{vmatrix} \begin{vmatrix} \omega_{X_1} \\ \omega_{Y_1} \\ \omega_i + \omega_{Z_1} \end{vmatrix} = \begin{vmatrix} J_{A_i}^x \omega_{X_1} \\ J_{A_i}^y \omega_{Y_1} \\ J_{A_i}^z (\omega_i + \omega_{Z_1}) \end{vmatrix} \quad (18)$$

Определим момент количества движения в системы:

$$\bar{L} = \bar{L}_C + \sum \bar{L}_i, \quad (19)$$

где $\bar{L}_C = I_C \bar{\omega}_C$ – кинетический момент корпуса относительно центра масс квадрокоптера;

$\bar{L}_i = I_i \bar{\Omega}_i = (I_{A_i} + m_i l^2) \bar{\Omega}_i$ – кинетический момент i -го ротора относительно центра масс квадрокоптера в системе координат $CX_i Y_i Z_i$.

Тензоры инерции корпуса I_C и i -го ротора I_i с учетом того, что главные оси инерции механической системы являются главными центральными осями инерции, равны:

$$I_C = \begin{vmatrix} J_C^{X_1} & 0 & 0 \\ 0 & J_C^{Y_1} & 0 \\ 0 & 0 & J_C^{Z_1} \end{vmatrix}; \quad I_i = \begin{vmatrix} J_{A_i}^x + m_i l^2 & 0 & 0 \\ 0 & J_{A_i}^y + m_i l^2 & 0 \\ 0 & 0 & J_{A_i}^z + m_i l^2 \end{vmatrix} \quad (20)$$

Тогда:

$$\bar{L}_C = \begin{vmatrix} J_C^{X_1} & 0 & 0 \\ 0 & J_C^{Y_1} & 0 \\ 0 & 0 & J_C^{Z_1} \end{vmatrix} \begin{vmatrix} \omega_{X_1} \\ \omega_{Y_1} \\ \omega_{Z_1} \end{vmatrix} = \begin{vmatrix} J_C^{X_1} \omega_{X_1} \\ J_C^{Y_1} \omega_{Y_1} \\ J_C^{Z_1} \omega_{Z_1} \end{vmatrix} \quad (21)$$

$$\bar{L}_i = \begin{vmatrix} J_{A_i}^x + m_i l^2 & 0 & 0 \\ 0 & J_{A_i}^y + m_i l^2 & 0 \\ 0 & 0 & J_{A_i}^z + m_i l^2 \end{vmatrix} \begin{vmatrix} \omega_{X_1} \\ \omega_{Y_1} \\ \omega_i + \omega_{Z_1} \end{vmatrix} = \begin{vmatrix} (J_{A_i}^x + m_i l^2) \omega_{X_1} \\ (J_{A_i}^y + m_i l^2) \omega_{Y_1} \\ (J_{A_i}^z + m_i l^2) (\omega_i + \omega_{Z_1}) \end{vmatrix} \quad (22)$$

С учетом (21), (22) выражение (19) будет иметь вид:

$$L = \begin{vmatrix} (J_C^{X_1} + \sum J_{A_i}^x + \sum m_i l^2) \omega_{X_1} \\ (J_C^{Y_1} + \sum J_{A_i}^y + \sum m_i l^2) \omega_{Y_1} \\ (J_C^{Z_1} + \sum J_{A_i}^z + \sum m_i l^2) \omega_{Z_1} + (\sum J_{A_i}^z + \sum m_i l^2) \omega_i \end{vmatrix} = \begin{vmatrix} J^{X_1} \omega_{X_1} \\ J^{Y_1} \omega_{Y_1} \\ J^{Z_1} \omega_{Z_1} + \sum J_i^z \omega_i \end{vmatrix}, \quad (23)$$

где $J^{X_1} = J_C^{X_1} + \sum J_{A_i}^x + \sum m_i l^2$, $J^{Y_1} = J_C^{Y_1} + \sum J_{A_i}^y + \sum m_i l^2$, $J^{Z_1} = J_C^{Z_1} + \sum J_{A_i}^z + \sum m_i l^2$, $\sum J_i^z = \sum J_{A_i}^z + \sum m_i l^2$ – приведенные осевые моменты инерции.

Для получения системы дифференциальных уравнений, описывающих вращение рассматриваемой системы, применим теорему об изменении кинетического момента механической системы [Павловский и др., 1990]:

$$\frac{d\bar{L}}{dt} = \frac{d\bar{L}}{dt} + (\bar{\omega}_C \times \bar{L}) = \sum \bar{M}_C^e \quad (24)$$

$$\frac{d\bar{L}}{dt} = \begin{vmatrix} J^{X_1} \dot{\omega}_{X_1} + \omega_{Y_1} \omega_{Z_1} (J_i^{Z_1} - J^{Y_1}) - \omega_{Y_1} \sum J_i^z \omega_i \\ J^{Y_1} \dot{\omega}_{Y_1} + \omega_{X_1} \omega_{Z_1} (J^{X_1} - J_i^{Z_1}) - \omega_{X_1} \sum J_i^z \omega_i \\ J^{Z_1} \dot{\omega}_{Z_1} + J_i^z \dot{\omega}_i + \omega_{X_1} \omega_{Y_1} (J^{Y_1} - J^{X_1}) \end{vmatrix} = \begin{vmatrix} M_{X_1}^e \\ M_{Y_1}^e \\ M_{Z_1}^e \end{vmatrix} \quad (25)$$

Тогда с учетом (12) и (25) система дифференциальных уравнений, описывающих движение квадрокоптера, может быть представлена в виде:

$$\begin{cases} m\ddot{X} = (\sin \psi \sin \phi + \cos \psi \cos \phi \sin \theta) \cdot \sum F_i \\ m\ddot{Y} = (\cos \phi \sin \psi \sin \theta - \cos \psi \sin \theta) \cdot \sum F_i \\ m\ddot{Z} = \cos \phi \cos \theta \cdot \sum F_i - mg \\ J^{X_1} \dot{\omega}_{X_1} + \omega_{Y_1} \omega_{Z_1} (J_i^{Z_1} - J^{Y_1}) + \omega_{Y_1} \sum J_i^z \omega_i = M_{X_1}^e \\ J^{Y_1} \dot{\omega}_{Y_1} + \omega_{X_1} \omega_{Z_1} (J^{X_1} - J_i^{Z_1}) - \omega_{X_1} \sum J_i^z \omega_i = M_{Y_1}^e \\ J^{Z_1} \dot{\omega}_{Z_1} + J_i^z \dot{\omega}_i + \omega_{X_1} \omega_{Y_1} (J^{Y_1} - J^{X_1}) = M_{Z_1}^e \end{cases} \quad (26)$$

Систему уравнений (26) необходимо решать совместно с кинематическими соотношениями, выражающими связь проекций угловой скорости аппарата и углов крена, тангажа и рыскания:

$$\begin{cases} \dot{\phi} = \omega_{X_1} - (\omega_{Z_1} \cos \phi - \omega_{Y_1} \sin \phi) \operatorname{ctg} \theta \\ \dot{\theta} = \omega_{Z_1} \sin \phi + \omega_{Y_1} \cos \phi \\ \dot{\psi} = \frac{1}{\cos \theta} (\omega_{Z_1} \cos \phi - \omega_{Y_1} \sin \phi) \end{cases} \quad (27)$$

Связь между угловой скоростью пропеллеров и величиной управляющих напряжений, поступающих на электродвигатели, запишем в виде [Емельянова и др., 2013а], [Емельянова и др., 2013б]:

$$\dot{\Omega}_i = \frac{-\frac{c_E c_M}{R} \eta N^2 \Omega_i - d \Omega_i^2 + \frac{c_M}{R} \eta N u_i(x, y, z, \phi, \theta, \psi)}{(J_p + \eta N^2 J_M)}, \quad i = 1-4 \quad (28)$$

где K_M – коэффициент пропорциональности, называемый постоянной момента электродвигателя; $N=\omega_i/\Omega_i$ – передаточное отношение редуктора, равное скорости электродвигателя ω_i , разделенное на скорость пропеллера Ω_i ; η -коэффициент эффективности (кпд), который связывает механическую энергию оси двигателя и пропеллера; J_p, J_M – момент инерции ротора вокруг оси винта и оси мотора соответственно; c_E, c_{Mi} – коэффициенты пропорциональности, называемые соответственно постоянной ЭДС двигателя и момента электродвигателя; d – аэродинамическая постоянная; R – активное сопротивление обмотки ротора.

Особый интерес представляет собой выбор стратегии управления, которая обеспечивает движение объекта по заданной траектории с заданной точностью и быстродействием и описывается функцией: $u_i = u_i(x, y, z, \varphi, \theta, \psi)$.

Пропорциональное управление по отклонению можно представить в виде:

$$u_i = u_i(x, y, z, \varphi, \theta, \psi) = u_{0i} \pm \Delta x k_x \pm \Delta y k_y \pm \Delta z k_z \pm \Delta \varphi k_\varphi \pm \Delta \theta k_\theta \pm \Delta \psi k_\psi, \quad (29)$$

где u_{0i} – постоянное напряжение питания; k_j – коэффициенты приращения; $\Delta x, \Delta y, \Delta z, \Delta \varphi, \Delta \theta, \Delta \psi$ – значения ошибок.

Для определения желаемой траектории движения и ориентации объекта в пространстве необходимо знать зависимость координат центра масс квадрокоптера от времени. Решение системы нелинейных дифференциальных уравнений (26), (27), (28), (29) осуществляется путем численного интегрирования, что позволяет найти зависимости $x(t), y(t), z(t), \varphi(t), \theta(t), \psi(t)$ и провести моделирование движения летающего робота.

Заключение

Предложена расчетная схема и математическая модель пространственного движения квадрокоптера, учитывающая гироскопические эффекты, вращающиеся винты, массогабаритные свойства четырех электроприводов, снабженных редуктором, кинематические связи, свойства электродвигателей, алгоритмы выработки управляющих воздействий.

В дальнейшем планируется создать адаптивный алгоритм, который заключается в использовании различных методик стабилизации по трем углам ψ, θ, φ для различных режимов и условий поле-

та. Решить задачу оптимального синтеза по критерию быстродействия при перемещении квадрокоптера из одной точки в другую. Выбрать области рациональных параметров, определяющих оптимальную работу приводов квадрокоптера в различных режимах для ПИД и логического регулятора.

СПИСОК ЛИТЕРАТУРЫ

[Емельянова и др., 2013а] Емельянова О.В., Попов Н. И., Яцун С. Ф. Моделирование движения квадроскопического летающего робота. Актуальные вопросы науки. Материалы VIII Международной научно-практической конференции. Москва, Спутник+. 2013, с. 6–8.

[Емельянова и др., 2013б] Емельянова О.В., Попов Н. И., Яцун С. Ф. Моделирование движения квадрокоптера в пространстве // Авиакосмические технологии (АКТ-2013). Труды XIV Всероссийской научно-технической конференции и школы молодых ученых, аспирантов и студентов. – Воронеж: ООО Фирма «Элист», 2013. – С.131–138.

[Попов и др., 2014а] Попов Н.И., Емельянова О.В., Яцун С.Ф., Савин А.И. Исследование колебаний квадрокоптера при внешних периодических воздействиях // Фундаментальные исследования. – 2014. – № 1. – С. 28–32.

[Попов и др., 2014б] Попов Н.И., Емельянова О.В. Динамические особенности мониторинга воздушных линий электропередачи с помощью квадрокоптера // Современные проблемы науки и образования. – 2014. – № 2; URL: <http://www.science-education.ru/116-12773> (дата обращения: 16.04.2014).

[Павловский и др., 1990] Павловский М.А., Акинфеев Л.Ю., Бойчук О.Ф. Теоретическая механика. Динамика. – К.: Выща шк., 1990. – 480 с.

[Bresciani, 2008] T.Bresciani. Modeling, identification and control of a quadrotor helicopter. Master's thesis, Department of Automatic control, Lund University, October 2008, p. 170.

[Tahar, 2011] M.Tahar, K.M.Zemalache, A. Omari Control of under-actuated X4-flyer using indegral Backstepping controller. Przegląd elektrotechniczny (Electrical review), ISSN 0033-2097, R.87 NR 10/2011, pages 251–256.

Список авторов

Абдулкадим Х. А.	243	Локтионова О.Г.	307
Афанасьева Т.В.	5	Малтугуева Г.С.	182
Ахмадеева И.Р.	209	Мескин П.И.	191
Аюпов М.М.	84	Миляков Д.А.	298
Берман А.Ф.	182	Минина Д.Ю.	77
Бодунков Н.Е.	275	Михайлов И.С.	40
Боровикова О.И.	209	Мошкин В.С.	173
Воронов С.В.	249	Наместников А.М.	146
Ворочаев А.В.	307	Наумов Г.С.	359
Ворочаева Л.Ю.	307	Невзорова О.А.	84
Гаврилова Т.А.	93	Николайчук О.А.	100
Грищенко М.А.	100	Новиков П.К.	289
Дементьев В.Е.	249, 256	Павловский В.Е.	315
Добрынин Д.А.	265	Палюх Б.В.	191
Дулин С.К.	109	Пантелеев М.Г.	324
Дулина Н.Г.	109	Петров А.В.	350
Емельянова О.В.	366	Пилипенко А.Ю.	289
Ефименко И.В.	118	Полковникова Н.А.	49
Ефимов С.В.	359	Поляков В.М.	31
Заварзин Д.В.	5	Попов Н.И.	366
Загорулько Г.Б.	200	Прокопов Е.Ю.	289
Загорулько Ю.А.	209	Прокопьев Н.А.	84
Зернов М.М.	13	Савин А.И.	366
Зо Мин Тайк	40	Савицкий А.В.	315
Иванов В.К.	191	Саенко И.Б.	219
Итыгилов В.Г.	200	Святкина М.Н.	58
Ким Н.В.	275, 283	Сидорова Е.А.	209
Клещев А.С.	128	Синюк В.Г.	31
Кобяков А.А.	137	Степанов Д.Н.	333
Ковалев С.М.	22	Субхангулов Р.А.	228
Кожунова О.С.	109	Сулейманов Д.Ш.	84
Колесов А.О.	146	Тарасов В.Б.	58
Кондратьев Д.С.	256	Тиминова И.А.	68
Корлякова М.О.	289	Тищенко И.П.	333
Котенко И.В.	219	Усталов Д.А.	237
Кохов В.В.	155	Филиппов А.А.	228
Кохтенко Н.В.	324	Хахалин Г.К.	164
Крылов И.Г.	283	Хачумов М.В.	342
Курбатов С.С.	164	Хитыков В.В.	350
Курейчик В.М.	49	Ходашинский И.А.	77
Куценко Д.А.	31	Хорошевский В.Ф.	118
Лапшин К.В.	137	Шалфеева Е.А.	128
Лебедев С.В.	324	Юрин А.Ю.	100, 182
Липатов А.А.	298	Яковлев К.С.	350
Лобзин А.П.	164	Ямщиков Ю.А.	137
Логинов М.И.	350	Ярушкина Н.Г.	173
Логинов С.С.	249	Яцун С.Ф.	359, 366

ABSTRACTS

METHODS OF JUSTIFICATION DEGREE CALCULATION IN THE ARGUMENTATION SYSTEMS

Oleg L. Morosin (*omorsik@gmail.com*)

Vadim N. Vagin (*vagin@appmat.ru*)

Moscow Power Engineering Institute

111250 Krasnokazarmennaja str., 14, Moscow, Russia

In this paper various methods of applying the degree of justification in argumentation systems is considered. Degrees of justification allow to model different reasoning tasks more efficiently and give a numerical estimate to the plausibility of a statement. In conclusion, the test example with justification degrees is given.

MODELING THE BEHAVIOR OF INTELLIGENT AGENT IN THE UNANTICIPATED INCREASE IN THE DURATION OF ITS DEDUCTIVE CYCLE UNDER TIGHT TIMELINES

Igor B. Fominych (*igborfomin@mail.ru*)

Moscow Power Engineering Institute

111250, Russia, Moscow, 14, Krasnokazarmennaja

Mikhail M. Vinkov (*vinkovmm@mail.ru*)

Russian Research Institute on Information and Computer Aided Technologies

129090, Russia, Moscow, 22, Stchepkina St.

The paper discusses the approach for improving rational agent operating in a hard real-time, to the emergence of unforeseen contingencies associated with its functioning. This approach is based on the concept of metacognition and implemented by means of the so-called active logics

DESIGN FEATURES OF BAYESIAN BELIEF NETWORKS FOR MODELING THE TEMPORAL REASONING

Andrew S. Zaharov (*auth1989@yandex.ru*)

Smolensk Branch of Moscow Power Engineering Institute

(Technical University), Smolensk, Russia

This paper proposes a mechanism for describing the temporal statements within the fuzzy Bayesian belief networks that allows the modeling of temporal reasoning based on them. Describes design features of the proposed hybrid model.

**METHODS OF CONSTRAINT INFERENCE AND HEURISTIC
SEARCH BASED ON THE MATRIX REPRESENTATION
OF FINITE PREDICATES**

Alexander A. Zuenko (*zuenko@iimm.kolasc.net.ru*)
Institute for Informatics and Mathematical Modelling, Kola Science Centre
of RAS, Russia, 184209, Murmansk Region, Apatity, Fersman St., 24A

This paper proposes methods of constraint inference and heuristic search for solving constraint satisfaction problems with the variables defined on finite domains. The developed methods use the matrix representation of finite predicates which allow to not convert constraints into binary relations and implement an effective reduction of the search space.

**PROCESSING OF SPECIAL KINDS OF CONSTRAINTS IN SOLVING
CONSTRAINT SATISFACTION PROBLEMS IN THE STRUCTURES
OF N-TUPLE ALGEBRA**

Alexander A. Zuenko (*zuenko@iimm.kolasc.net.ru*)
Institute for Informatics and Mathematical Modelling, Kola Science Centre of
RAS, Russia, 184209, Murmansk Region, Apatity, Fersman St., 24A

The paper proposes modification of basic methods of constraint inference, previously developed for solving constraint satisfaction problems with finite domains. The methods use matrix structures of n-tuple algebra. The modification applies additional rules for search space reduction, which are determined by subject domain specifics and significantly accelerate the problem solving.

**ABOUT TEMPORAL REASONING COMPONENT BASED ON POINT-
INTERVAL LOGIC FOR INTELLIGENT DECISION SUPPORT
SYSTEMS OF REAL TIME**

I.E. Kurilenko (*kurilenkoiy@mpei.ru*)
National Research University “Moscow Power Engineering Institute”,
Krasnokazarmennaya 14, Moscow, 111250 Russia

D.A. Shornikova (*daria-chan@mail.ru*)
National Research University “Moscow Power Engineering Institute”,
Krasnokazarmennaya 14, Moscow, 111250 Russia

The paper presents inference algorithms for point-interval temporal logic. Inference mechanism based on converting temporal information from temporal relations to point graph is shown. The paper demonstrates PIL inference engine for intelligent decision support system of a real time.

ABOUT SOLUTIONS OF A PRODUCTION-LOGICAL EQUATION ON A BOOLEAN LATTICE

S.D. Makhortov (*sd@expert.vrn.ru*)
Voronezh State University

I.Y. Ivanov (*hour1scorp@gmail.com*)
Voronezh State University

The class of production-logical equations on a finite boolean lattice is introduced. The questions of solvability of such equations are considered. The concepts of a canonical relation on a boolean lattice and a structural layering of a canonical relation are introduced that allows to reduce a finding of an approximate solution of an equation to the directed graph's nodes traversal problem. The results may be applied to a logical inference optimization and corresponding knowledge bases verification.

METHOD OF ACCELERATED LOGICAL INFERENCE FOR FORMAL VERIFICATION PROBLEM SOLVING

Gennadiy A. Chistyakov (*gennadiychistyakov@gmail.com*)
Vasiliy Yu. Meltsov (*meltsov69@mail.ru*)
Vyatka State University

Russia, Kirovskaya obl., Kirov, Moskovskaya 36, 610000

Paper contains a statement of formal verification logical problem and briefly description of the specialized inference method that is used to accomplish this task. The method takes into account the structures of premises in the knowledge base that allows to search the possible solutions to the most efficient manner.

EVALUATION OF HYBRID OUTPUTS IN PRODUCTION SYSTEMS

Sergey Y. Soloviev (*soloviev@glossary.ru*)
Moscow State University
CM&C faculty, 2 educational corps, MSU, 119992

In this paper we consider hybrid inferences built one product system involving rules of other production system. Focuses on a method for calculating estimates of originality hybrid conclusions.

CONCEPTION OF MULTIAGENT TEXT ANALYSIS SYSTEM FOR ONTOLOGY POPULATION

Anokhin S.A. (*saanokhin@gmail.com*)

Garanina N.O. (*garanina@iis.nsk.su*)

Sidorova E.A. (*lena@iis.nsk.su*)

A.P. Ershov Institute of Informatics Systems,
6, Acad. Lavrentjev pr., Novosibirsk 630090, Russia

The paper is focused on the conception of multiagent text analysis system for domain-specific ontology population. A formal knowledge model is presented. We describe multiagent model and multiagent text analysis algorithm based on interaction of informational agents and rule-agents. Overview of the proposed system's architecture is given.

APPROACHES TO THE ONTOLOGY ALIGNMENT AND IDENTITY RESOLUTION PROBLEMS WHEN ENRICHING THE CONTENT OF SCIENTIFIC KNOWLEDGE BASES WITH THE LINKED OPEN DATA

Zinaida V. Apanovich (*apanovich@iis.nsk.su*)

Alexander G. Marchuk (*mag@iis.nsk.su*)

A.P. Ershov INSTITUTE OF INFORMATICS SYSTEMS, Russian Academy
of Sciences, Siberian Branch
6, Acad. Lavrentjev pr., Novosibirsk 630090, Russia

This paper describes approaches to the vocabulary normalization and identity resolution problems arising during the use of the LOD datasets to enrich the content of scientific knowledge bases. The dataset of the Open Archive of the Russian Academy of Sciences, and several bibliographic datasets are used as test examples.

FORMAL MODEL OF KAZAKH NOMINAL

B.Zh. Yergesh (*b.yergesh@gmail.com*)

A.S. Mukanova (*asel-ms@bk.ru*)

G.T. Bekmanova (*gulmira-r@yandex.ru*)

A.A. Sharipbay (*sharalt@mail.ru*)

L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
010000 Kazakhstan, Astana, 2, Mirzoyana str., teaching and administrative
(main) building of the ENU

This paper explains how semantic hypergraphs are used to construct ontological models of nominal in the Kazakh language.

RESTORATION OF GRAMMATIC ELLIPSES IN SYNTACTIC ANALYSIS

Tatiana Yu. Kobzareva (*t.kobzareva@gmail.com*)
Michael E. Epifanov (*xeme@rambler.ru*)
Delir G. Lakhuti (*delir1@yandex.ru*)
Russian State University for Humanities, Moscow

We consider the problem of grammatical ellipses restoration in automatic syntactic analysis of Russian sentence. Are discussed grammatical conditions for the occurrence of four most probable types of ellipses enabling the location of their antecedents.

NETWORK OF NATURAL HIERARCHIES OF TERMS AN ANALYSIS OF SCIENTIFIC TEXTS

D.V. Lande (*dwlande@gmail.com*)
Institute for Information Recording NAS of Ukraine, Ukraine, Kiev
A.A. Snarskii (*asnarskii@gmail.com*)
NTUU “Kiev Polytechnic Institute” Ukraine, Kiev
E. V. Yagunova (*iagounova.elena@gmail.com*)
Saint-Petersburg State University Russian Federation, St.-Petersburg

The technique of building of networks of hierarchies of terms based on the analysis of scientific texts is offered. The technique is based on the methodology of horizontal visibility graphs for the terms – of individual words, bigrams and trigrams, as well as an inclusion relations between the terms. Constructed and investigated language network, formed on the basis of automated processing of full papers conference OSTIS-2014.

RUTHES THESAURUS: PUBLICATION OF THE FIRST VERSION

N.V. Loukachevitch (*louk_nat@mail.ru*)
B.V. Dobrov (*dobrov_bv@mail.ru*)
Lomonosov Moscow State University

The paper presents RuThes-lite, a publicly available version of RuThes linguistic ontology, which has been developed for more than fifteen years and is intended for automatic document processing. RuThes-lite was generated from RuThes on the basis of the most frequent words in a contemporary news collection. In this paper we describe the structure of RuThes-lite and ways of access to its data.

CONCEPTUAL SCHEME FOR TEXT CLASSIFICATION SYSTEM USING JSM METHOD

Nickolay D. Lyfenko (*lyfenkoNick@ya.ru*)
Russian State University for the Humanities
Miusskaya sq. 6, Moscow, GSP-3, 125993, Russia

The paper describes an application of classification algorithms to the text categorization problem. Author proposes a conceptual scheme for an automatic text categorization system. This system must operate with various text representation models and data mining methods. The novelty of this system consists in advanced implementation of JSM method for automatic hypothesis generation – an original logical-combinatorial technology of data mining.

AUTOMATIC TERM EXTRACTION FOR THE TASK OF SENTIMENT CLASSIFICATION FOR DYNAMICALLY UPDATED TEXT COLLECTIONS

Y. Rubtsova (*yu.rubtsova@gmail.com*)
The A.P. Ershov Institute of Informatics Systems (IIS),
Siberian Branch of the Russian Academy of Sciences, Novosibirsk

The paper presents an approach to automatic terms extraction from micro blog text collection for constructing continuously updated vocabulary. The prepared dictionary is used for the task of sentiment classification of text collections into three classes (positive, neutral, negative). The results of sentiment classification and accuracy comparison between methods based on different weighting schemes are also described. In addition the computational complexity of generating representations for time-dependent text collections depending on weighting schemes is discussed.

MORPHEME-BASED SPEECH RECOGNITION SYSTEM FOR THE TATAR LANGUAGE

Dzh.Sh. Suleymanov (*dvd.slt@gmail.com*)
A.F. Khusainov (*khusainov.aidar@gmail.com*)
Institute of Applied Semiotics, Tatarstan Academy of Sciences, Kazan, Russia
Kazan (Volga region) Federal University, Kazan, Russia

This paper describes speech recognition system for the Tatar language. Automatic phonetic transcription system for Tatar texts has been developed and implemented; this system based on acoustic rules found for the Tatar language. Morpheme-based approach used to build language model for the Tatar. Proposed system showed 75% recognition accuracy in test subcorpus

ARTIFICIAL INTELLIGENCE: SCIENCE MAP OF THE DOMAIN IN CAI PROCEEDINGS

V.F. Khoroshevsky (*khor@ccas.ru*)
Dorodnitsyn Computing Centre RAS, Moscow
I.V. Efimenko (*iefimenko@hse.ru*)
Higher School of Economics, Moscow

The paper represents the research aimed at producing science maps of Artificial Intelligence based on bibliometric analysis of the papers in CAI proceedings (1996–2012). The corpora were processed using ontology based Information Extraction (OBIE), and a variety of clusterization methods. Analysis of the received science maps was carried out with time series, tag clouds and the ‘shadow groups’ method.

INTELLECTUAL COGNITIVE SYSTEM OF REAL TIME FOR CONTROL OF SAFETY OF WORK OF NUCLEAR POWER PLANTS OF CONCERN “ROSENERGOATOM”

A.A. Bashlykov (*Aleks-ltd@mail.ru*)
M.A. Lygin (*Lygin@transenergostroy.ru*)
National research university MPEI; “ORGENERGOSTROY LTD.”
Krasnokazarmennaya St., 14, Moscow, 111250; Derbenevskaya Embankment,
7/10 Moscow, 115114;

In article questions of application of means of cognitive graphics for expeditious support of decision-making by operation personnel of the Crisis center of Rosenergoatom concern are considered at the solution of problems of expeditious monitoring and control of implementation of the schedule of execution of loading by power units of nuclear power plants in the generalized parameters.

COGNITIVE SEMIOTICS: SEMIOSIS AND GESTALTS

Yuri R. Valkman (*yur@valkman.kiev.ua*)
International Research Training Center for Information Technologies and
Systems National Academy of Sciences and the Ministry for Education and
Science of Ukraine,
Glushkova str. 40, Kyiv-187, 03680, UKRAINE

It is demonstrated that gestalts are basic sources as well as semioses are basic processes in cognitive semiotics operations. The properties of gestalts are studied. The classification of semiosis schemes is considered. By means of these schemes an interpretation of signs is made. The principles of building the semiosis schemes are examined to construct the formal apparatus for further development of computer technologies of imagery thinking.

VISUAL IMAGES-AND-SYMBOLS IN ARTIFICIAL SYSTEMS

Boris A. Kobrinskiy (*bakob@pedklin.ru*)

Pirogov Russian National Research Medical University (Moscow)
Russia, 125412, Taldomskaya str, 2, Research Clinical Institute for Pediatrics
of Pirogov Russian National Research Medical University

The paper describes the principles for the transition to visual images in intelligent systems. The application of various types of images series (fuzzy, temporary and other) knowledge bases on the example of medical diagnostic systems.

THE APPROACH TO ADVANCE OF COGNITIVE RESEARCH METODOLOGICS IN VIRTUAL COMMUNICATION

J.M. Kuznetsova (*kuzjum@yandex.ru*)

N.V. Chudova (*nchudova@gmail.com*)

Institute for Systems Analysis of Russian Academy of Sciences (ISA RAS);
117312, Moscow, pr. 60-letiya Oktyabrya, 9

This paper considers the applicability of psycholinguistics and AI for reconstruction the Model of the World by analyzing online communication.

ON THE CONDITIONALITY OF INDIVIDUAL COGNITIVE STYLE FORMATION

A.V. Pavlov (*pavlov@phoi.ifmo.ru*)

K.A. Isakov

National Research University of Informational Technologies,
Mechanics and Optics, Saint Petersburg

Formation of individual cognitive style is considered in the framework of logic with exception based on algebra of Fourier dual operations. The connection of cognitive style with described by dissipative factor individual physical properties and experience of cognitive agent has been revealed. A number of cognitive dissonance reducing scenarios were modeled, mechanism of hysteresis initiation has been demonstrated.

CONFLICT RESOLUTION METHODS IN HYBRID INTELLIGENT SYSTEMS

Vitaly V. Smirnov (*vitaly_smirnov@mail.ru*)

Delis Archive; 107023, Moscow, Elektrozavodskaya st., 21-41-XXII

Igor. B. Phominikh (*igborfomin@mail.ru*)

National Research University "Moscow Power Engineering Institute";
111250, Moscow, Krasnokazarmennaya st., 14

The paper discusses methods to resolve conflicts in hybrid intelligent systems, shows the applicability of these methods, depending on the perception levels of internal and external environments, presents several common points of view on the mental processes of perception.

SIMULATION OF CONDUCT OF MEMBERS OF VIRTUAL COMMUNITY AT PORTAL ASPIRANTURA.SPB

V.I. Tishchenko (*tishchenkovictor@gmail.ru*)

A.L. Prochko (*alexei@prochko.ru*)

Institute for Systems Analysis of RAS, Moscow

It is demonstrated the use of the principles of the theory of complex systems for modeling virtual communications. The study of the scientific community on the portal aspirantura.spb.ru obtained graphical representation (graph) of social network participants. It is shown that the use of the algorithm reveals Luvinskogo network of virtual communities, whose members are characterized by varying degrees of activity and betweenness centrality (connectedness).

THE INSPIRED BY NATURAL SYSTEMS ALGORITHM FOR SOLVING PROBLEMS OF DESIGN AND OPTIMIZATION

V.V. Bova (*vvbova@yandex.ru*)

V.V. Kureichik (*vkur@tgn.sfedu.ru*)

VI.VI. Kureichik (*vkur@tgn.sfedu.ru*)

Southern Federal University, Taganrog

The algorithms inspired by natural systems for solving design and optimization are described in this paper. These algorithms allow to parallelize the search process and eliminating the problem of partially pre-convergence, as well as receive optimal sets and suboptimal solutions in polynomial time. Time complexity of algorithms in the best case $\sim O(n \log n)$, in the worst case $- O(n^2)$.

THE DECISION OF A PROBLEM OF PLACEMENT OF ELECTRONIC ELEMENTS ON THE BASIS OF FUZZY GENETIC ALGORITHM

Leonid A. Gladkov (*leo_gladkov@mail.ru*)
Nadezhda V. Gladkova (*nadyusha.gladkova77@mail.ru*)
Sergey N. Leiba (*leo@tgn.sfedu.ru*)
Southern Federal University,
347928, Russia, Taganrog, 44 Nekrasovskiy lane

In work the problem of placement of electronic elements on a patchboard is considered. The placement problem is one of problems of designing of topology and it also is a NP-full problem. In article the approach to the decision of a task in view on the basis of integration of genetic algorithms of search and fuzzy models of management in parameters of algorithm is offered. The description of structure of the used fuzzy logic controller is given. The short description of the program and the results of its testing confirming efficiency of the offered method are resulted.

USING A SELF-ADAPTIVE POPULATION SIZE IN MULTI- OBJECTIVE GENETIC ALGORITHMS

Pavel V. Kazakov (*pvk_mail@list.ru*)
Bryansk state technical university
Russia, 241035, Bryansk, boulevard of 50-year of October, 7

The new manner for improving of the multi-objective genetic algorithms is considered. It based on the new model for a population size control. This approach is universal and may be implemented for any multi-objective evolutionary algorithms. The comparative performance analysis of the using suggested manner for solving different multi-objective problems are cited.

MODEL OF INTERACTION BETWEEN LEARNING AND EVOLUTION: MECHANISMS OF INTERACTION

Vladimir G. Red'ko (*vgredko@gmail.com*)
Scientific Research Institute for System Analysis, Russian Academy of Sciences
Vavilova Street, 44/2, Moscow, 119333, Russia

The model of interaction between learning and evolution is designed and investigated. Following mechanisms of this interaction are investigated. 1) The mechanism of the genetic assimilation of individually acquired features. 2) The hiding effect: strong learning can inhibit evolutionary optimization. 3) The mechanism of influence of the learning load on the interaction.

SYMMETRIC KEY CRYPTOGRAPHY USING GENETIC ALGORITHM

Vadim Stefanuk (*stefanuk@iitp.ru*)

Institute for Information Transmission Problems of RAS, Moscow

Amanie Alhussain (*amanie-alhussain@hotmail.com*)

People's Friendship University of Russia, Moscow

Unlike the well known use of Genetic Algorithms within encryption keys in the present paper it is proposed to include GA directly into transmission path intended for the encrypted information, where this algorithm improves the properties of the Random Number Generators. The usefulness of our approach is illustrated with some examples. Moreover, the paper demonstrates an implemented cryptography system, where the principles described in the paper have been applied.

AUTOMATED NEURAL NETWORK MODELS FORMING FOR POPULATION MORBIDITY LEVEL PREDICTION

D.I. Khritonenko (*hdmitry.91@mail.ru*)

E.S. Semenkin (*eugenesemenkin@yandex.ru*)

E.V. Sugak (*sugak@mail.ru*)

E.N. Potilicina (*leonova_en@mail.ru*)

Siberian state aerospace university named after academician M.F. Reshetnev,
Krasnoyarsk

The automated forming of artificial neural networks with self-configured evolutionary algorithms is considered. The obtained algorithm solves the problem of predicting the population morbidity level depending on the level of atmosphere contamination. A series of tests and comparisons of different variants of the proposed algorithm has been held.

TIME SERIES' SUBSTANTIVE INTERPRETATION BASED ON GENETIC ALGORITHM.

L.S. Shigabutdinova (*liliya.scherbakova@gmail.com*)

Ulyanovsk State Technical University

432027, Ulyanovsk reg., Ulyanovsk city, Severnyi Venets st., 32

This article is devoted to the problem of time series' substantive interpretation. Readers are invited to software product that allows to recognize the trends of the time series, as well as to get a linguistic assessment of its fragments. System functioning includes two stages: the transformation of the initial time series using a genetic algorithm, and ontological analysis of the resulting series to obtain expert linguistic interpretation of its parts.

GAME THEORETIC ALGORITHMS OF WIRELESS NETWORK FORMATION

Nikolay Bazenkov (*n.bazenkov@gmail.com*)
Trapeznikov Institute of Control Sciences
65 Profsoyuznaya street, Moscow 117997, Russia

We study the problem of efficient topology formation for wireless ad hoc networks. The problem is to form a connected network with minimal total power of wireless transmitters. The algorithms employing novel game-theoretic method called double best response are considered and their efficiency and computational complexity are investigated.

MULTI-LATERAL GOAL-RESOURCE NETWORKS IN THE MODELING OF COMMUNICATION BETWEEN COGNITIVE AGENTS

V.S. Dyundyukov (*vsd89@yandex.ru*)
Bauman Moscow State Technical University
V.B. Tarassov (*vbulbov@yahoo.com*)
Bauman Moscow State Technical University

A new weighted graph model called goal-resource network is introduced in order to take into account agent types and their behavior while resource exchanging and multi-agent system formation. Primarily, some important prerequisites of transition from resource networks to goal-resource networks related to agent architecture and various agent types are discussed. A formal definition of goal-resource network is given. Main types of interactions between agents are represented by bilateral weighted graphs. Various situations of resource exchanges are considered for agents of different types.

DYNAMIC GAME WITH LINEAR UTILITY FUNCTION ON COGNITIVE MAP WITH LIMITED AND UNLIMITED RESOURCE

Sergei Kulivets (*skulivec@yandex.ru*)
V.A. Trapeznikov Institute of Control Sciences RAS;
65 Profsoyuznaya street, Moscow 117997, Russia;

We considered two game-theoretic models of interaction between agents during the time for linear cognitive map. The first model describes the interaction of agents with constant restrictions on their effects at each time point. The second model considers the situation when the agent control factor impact is a limited resource throughout the game. We found a solution in the form of Nash-equilibrium strategies

for both models. Considered models can be applied in inconsistent representation case without significant changes.

A FORMAL MODEL OF PROACTIVE ITERATIVE PLANNING FOR REAL-TIME INTELLIGENT AGENTS

Michael G. Panteleyev (*MPanteleyev@gmail.com*)

Saint-Petersburg State Electrotechnical University "LETI" Prof. Popova st. 5,
Saint-Petersburg, 197376, Russia *MPanteleyev@gmail.com*

The problem of planning actions of intelligent agents in dynamic multi-agent worlds is considered. The formal models of knowledge representation and processing necessary for implementation of proactive iterative planning are discussed. An approach and models are illustrated by examples of situations from virtual football.

THE MULTI-AGENT SYSTEMS MODEL OF AN INFORMATIONAL AND EDUCATIONAL SPACE

Y.F. Telnov

The Moscow State University of Economics, Statistics and Informatics (MESI),
Moscow

The report proposes a realization of multi-agent systems model of informational and educational space (MAS IES) of higher education institutions, which provides integration of the heterogeneous knowledge sources and services in distributed computing network for the realization of collective access to different categories of participants in the scientific and educational processes. The report reveals the features of the implementation of multi-agent systems model, linking the intelligent agents of the agents of scientific and educational activities with a repository of scientific and educational objects and services.

MULTI-AGENT SYSTEM FOR SOLVING PROBLEMS OF PURPOSEFUL BEHAVIOR

V.M. Trembach (*trembach@yandex.ru*)

Moscow Aviation Institute (National Research University), "MAI"
125871, Moscow, Volokolamskoe sh., 4

The paper describes a multi-agent system for solving problems of purposeful behavior. An integrated approach to represent knowledge is used, that allows you to shape and implement plans to achieve goals. And the examples of the implementation of intelligent agents are also considered.

SELF-ORGANIZATION MODEL IN THE AGENT MESAHE TRANSMISSIONING SYSTEMS

Anatoly M. Fedotov (*fedotov@sbras.ru*)
Institute of Computation Technologies SB RAS, Novosibirsk
Sergey G. Lomakin (*sir_ejik@mail.ru*)
Novosibirsk State University, Novosibirsk

This thesis describes how predominant opinion of majority could be changed completely by the small number of randomly distributed agents, who promote their opinion and are not vulnerable to the influence of other opinions. Existence of the critical moment, after which the promoted opinion is accepted by the community is shown in the thesis as well.

IMPLEMENTATION OF CASE-BASED REASONING MODULE FOR INTELLIGENT DECISION SUPPORT SYSTEM

Roman V. Alekhin (*r.alekhin@gmail.com*),
Pavel R. Varshavskiy (*VarshavskyPR@mpei.ru*),
National Research University "MPEI", Moscow

The paper addresses the issues of implementation of the case-based reasoning module for intelligent decision support system (IDSS). The method of case-based search for a solution with the use of structure mapping theory and ontology domain is proposed. The possibility of solving the problem of expert diagnosis of a complex object using the prototype of IDSS's case-based reasoning module for one of the power subsystems is considered.

THE SITUATIONAL HUMAN-MACHINE INTERFACE FOR INFORMATION SYSTEMS OF INTELLECTUAL SUPPORT OF DECISION-MAKING

A.A. Bashlykov (*Aleks-ltd@mail.ru*)
National research university MPEI;
Krasnokazarmennaya St., 14, Moscow, 111250;

Results of the analysis of operator interfaces are given in article for intellectual support of adoption of operating decisions in human-machine control systems of difficult objects. The concept of the situational human-machine interface based on methods of cognitive graphics is entered and the principles of its construction are described. difficult objects. The example of the organization of such situational interface for dispatchers of pipeline systems is given.

VERIFICATION OF PROCESS MODELS ON THE BASIS OF TEMPORAL PETRI NETS

Alexander P. Eremeev (*eremeev@appmat.ru*)

Yury I. Korolev (*korolevyu@gmail.com*)

National Research University "Moscow Power Engineering Institute"
Krasnokazarmennaya 14, Moscow, 111250 Russia

The verification of Petri nets temporal subclass, proposed for use in the dynamic intelligent systems of the type of real time intelligent decision support systems, for modeling of processes in complex objects are considered. The graph analysis tools, reflecting the changes in such networks, and the possibility of applying for verification of model checking method are described.

THE OPERATIVE ASSIGNMENT OF THE PURPOSE IN THE ANTHROPOCENTRAL OBJECTS WITH POSITIONS OF THE MODEL "STAGE"

C.YU. Zheltov (*zhl@gosnias.ru*)

B.E. Fedunov (*boris_fed@gosnias.ru*)

FGUP GOSNIAS (State Research Institute of Aviation Systems)
Moscow, Russia

Operative assignment of the purpose is "the primary phase of the management" into anthropocentric objects (Anthr/objects). In the ideas of the models "Stage" it is an operative changing to standard situation (TS) of the current session of the functioning of the Anthr/object. The structure of the algorithm of "The operative assignment of the purpose": i) a priori given the sequence of the TS-s (the multitude named "A"), providing the decision of the general problem of the coming session, ii) a priori given multitude of the TS-s (the multitude named "B"), which can appear (as the obstacles) in the coming session, iii) the matrix of the knowledge's with the terms of the linguistic variables for every couple of the TS from the multitude "A" and the multitude "B"; iv) the choice of the priority TS with the help of the matrix of the knowledge.

HYBRID INTELLIGENT SYSTEMS: COORDINATION AND CONSISTENCY

Igor A. Kirikov (*baltbipiran@mail.ru*)

Kaliningrad branch of the Institute of Informatics Problems of the Russian Academy of Sciences, Kaliningrad, Gostinaya str., 5

Alexander V. Kolesnikov (*avkolesnikov@yandex.ru*)

Immanuel Kant Baltic Federal University, Kaliningrad, Nevskogo str., 14
Kaliningrad branch of the Institute of Informatics Problems of the Russian

Academy of Sciences, Kaliningrad, Gostinaya str., 5

Sergey V. Listopad (*ser-list-post@yandex.ru*)

Kaliningrad branch of the Institute of Informatics Problems of the Russian Academy of Sciences, Kaliningrad, Gostinaya str., 5

Sergey A. Soldatov (*baltbipiran@mail.ru*)

“HIMAS +”, Kaliningrad, Sovetskiy av., 1

The work deals with modeling of heterogeneity of decision support systems based on coordination and consistency. The proposed models have been tested on the complex tasks of production planning and logistics.

COGNITIVE ARCHITECTURE OF DECISION-MAKING SUPPORT SYSTEMS: THE SEMIOTIC APPROACH

Alexander A. Kulinich (*kulinich@ipu.ru*)

V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences.
Profsoyuznaya 65, Moscow, Russian Federation

Cognitive architecture of decision-making support systems constructed on the basis of intelligence model as an organization forms of the mental experience and by principles of semiotic system is considered.

PECULIARITIES OF BUILDING HIERARCHICAL CRITERIA SYSTEM FOR EVALUATION OF TECHNICAL COMPLEX

V.N. Lobanov (*Fisher_1980@mail.ru*)

M.A. Kartsev Scientific Research Institute for Computer Complexes Inc., Moscow

A.B. Petrovsky (*pab@isa.ru*)

Institute for Systems Analysis, Russian Academy of Sciences, Moscow

The paper considers the main features of aggregating initial attributes, which are described the technical complex, in the process of building a hierarchical system of composite criteria and an integral indicator for multiple criteria evaluation and selection of the complex.

**KNOWLEDGE ELICITATION
ABOUT GROUPS OF TIME SERIES
FOR FORECASTING PROBLEM**

Irina G. Perfilieva (*irina.perfilieva@osu.cz*)

University of Ostrava, 30. dubna 22 701 03 Ostrava 1, Czech Republic

Anton A. Romanov (*romanov73@gmail.com*)

Ulyanovsk state technical university, 32 Severny Venetz str., 432027
Ulyanovsk, Russia

Paper describes an approach to forecasting time series groups. The method of F-transform of the highest order is used for analyze the rate of change of the piecewise trend time series. We propose to classify the time series by degree of similarity of such changes. Using neural network prediction methods for only one representative from the group will allow to predict the behavior of other similar time series.

**MULTICRITERIA APPROACH IN RESOURCE ALLOCATION
PROBLEM**

Eugene Y. Roitman

Geoprom, (*reja41@yandex.ru*)

Gregory V. Royzenson

Institute for Systems Analysis, Russian Academy of Sciences,
Prospect 60 Let Otyabrya, 9, Moscow 117312, Russia, (*rgv@isa.ru*)

The solution of a problem of optimum allocation of resources (production schemes) for the nonferrous metallurgy factories is in-process observed. Methods are involved in the presented approach multicriteria linear programming and methods of the group verbal decision analysis. Formulation of the problem is proposed for multiobjective optimization forming a Pareto optimal set of alternatives and the use of the group ranking method for selecting the best multicriteria alternatives. Formed a set of evaluation criteria for selecting a scheme of production. The instance of the solution of a practical problem of optimum allocation of resources between the nonferrous metallurgy factories is presented. Directions of use of the presented approach for the solution of variety of actual problems in nonferrous metallurgy are offered.

**INTELLIGENT SUBSYSTEM FOR FAULT-TOLERANT
UNCONDITIONAL DIAGNOSTIC TESTS CONSTRUCTION
AND DECISION-MAKING ON ITS BASE**

Anna E. Yankovskaya (*ayyankov@gmail.com*)
Tomsk State University of Architecture and Building 2, Solyanaya Square,
634003, Tomsk National Research Tomsk State University
36, Lenin Ave., 634050, Tomsk
Tomsk State University of Control Systems and Radioelectronics
40, Lenin Ave., 634050, Tomsk
Alexander I. Gedike (*gai@tsuab.ru*)
Tomsk State University of Architecture and Building
2, Solyanaya Square, 634003, Tomsk

The matrix model of data and knowledge representation is briefly described. The mathematical foundation of fault-tolerant irredundant diagnostic tests construction are expounded. The architecture and principles of software implementation of intelligent subsystem for fault-tolerant irredundant diagnostic tests construction and decision-making on its basis are given.

**DEVELOPMENT OF INTELLIGENT SERVICES WITH VIRTUAL
REALITY**

Valeria V. Gribova (*gribova@iacp.dvo.ru*),
Leonid A. Fedorischev (*fleo@mail.ru*)
Institute of Automation and Control Processes
Radio st., 5, Vladivostok, Russia, 690041

The paper is devoted to description of the tool, the platform, and the technology for creation a new class of software: intelligent systems with virtual reality.

**MODEL OF MANIFESTATION OF EMOTIONS IN NATURAL
SPEECH**

N.N. Filatova (*nfilatova99@mail.ru*)
D.M. Haneev (*t_ran@mail.ru*)
I.A. Rebrun (*igar_rebrun@mail.ru*)
K.V. Sidorov (*bmisidorov@mail.ru*)
Tver state technical university

The task of the formalized description of influence of change of an emotional condition (registered on the basis of objective data) on process of formation of speech messages on the basis of the analysis of the correlated data from two sources of signals (a brain and organs of articulation) is considered.

**SOME ASPECTS OF INTEGRATED EXPERT SYSTEMS
CONSTRUCTION CONTROL WITH USE OF AUTOMATED
PLANNING**

Galina V. Rybina, Yuri M. Blokhin
(*galina@ailab.mephi.ru*);
National Research Nuclear University MEPhI
(Moscow Engineering Physics Institute)
115409, Kashirskoe sh. 31., Moscow, Russian Federation

Intellectualization of integrated expert systems construction processes based on task-oriented methodology and AT-TECHNOLOGY workbench are reviewed. Automated planning methods applied for expert system prototype architecture model generation with help of intelligent planner experience is described.

**THE APPROACH
TO THE TEMPORAL REASONING IMPLEMENTATION
IN DYNAMIC INTEGRATED EXPERT SYSTEMS**

Galina V. Rybina (*galina@ailab.mephi.ru*)
Aleksey V. Mozgachev (*avmozgachev@mephi.ru*)
National Research Nuclear University MEPhI
(Moscow Engineering Physics Institute)
Kashirskoye shosse 31, Moscow, 115409, Russian Federation

The scientific and technological problems of dynamic integrated expert systems creating and approaches to their solution are discussed. The development of task-oriented methodology for the integrated expert systems creating is considered in the context of creating dynamic integrated expert. Particular attention is paid to the temporal knowledge representation and processing.

**INTELLIGENT SYSTEM FOR FORMING PROJECT TEAMS BASED
ON COMPETENCIES.**

D.Y. Tikhomirova (*darya.tikhomirova@gmail.com*)
V.V. Sherkunov (*v.sherkunov@gmail.com*)
Ulyanovsk State Technical University
432027, Ulyanovsk reg., Ulyanovsk city, Severnyi Venets st., 32

This paper describes the ontological methods of search experts in the scientific organizations based on their competences. The method of formation of a conceptual model of the qualification reference manual by semantic analysis of the text

document in which it contains is provided. The data transformation method about education of the candidate for the list of its competences is considered. The genetic algorithm used to obtain the optimal distribution of experts is provided. The short review of developed intellectual system is given.

BASICS OF INTELLIGENT SYSTEM CREATION OF PROBABLE INFORMATION SECURITY ATTACKER SEARCH

Anna E. Yankovskaya (*ayyankov@gmail.com*)

Tomsk State University of Architecture and Building
2, Solyanaya Square, 634003, Tomsk

National Research Tomsk State University

36, Lenin Ave., 634050, Tomsk

Tomsk State University of Control Systems and Radioelectronics

40, Lenin Ave., 634050, Tomsk

Valentina G. Mironova (*mvg@security.tomsk.ru*)

Tomsk State University of Control Systems and Radioelectronics

40, Lenin Ave., 634050, Tomsk

Aleksander A. Shelupanov (*saa@keva.tusur.ru*)

Tomsk State University of Control Systems and Radioelectronics

40, Lenin Ave., 634050, Tomsk

The basics of intelligent system creation of probable information security attacker search are proposed. Intelligent system intended for subjects revealing which can cause a negative influence on information. This system is based on the matrix way of data and knowledge representation and test methods of pattern recognition.

VOTING PROCEDURE IN HOMOGENEOUS COLLECTIVES OF ROBOTS

Valery Karpov (*vkarpov@hse.ru*)

MIEM, National Research University Higher School of Economics,
Moscow

In work solutions of problems of definition of the leader and redistribution of roles in homogeneous group of robots are proposed. It is shown that using exclusively local interaction, transition from a swarm to collective of robots with the hierarchical organization is possible. At the heart of procedure of a choice of the leader the algorithm of local revoting lies, and redistribution of roles can be carried out by a wave method.

MODEL OF AGENTS (ROBOTS) COOPERATION

Alexander A. Kulinich (*kulinich@ipu.ru*)
V.A. Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences.
Profsoyuznaya 65, Moscow, Russian Federation

The model of agents (robots) cooperation in the dynamic situations, based on criteria of utility of the agent and cognitive dissonance of their relations is investigated.

SITUATION CONTROL AS THE BASIS OF ADAPTIVE CONTROL SYSTEM OF WALKING ROBOT

Vasily A. Lutsky (*Lu10001@yandex.ru*)
Moscow institute of electronics and mathematics of National Research
University "Higher school of economics" (MIEM HSE)
109028, Moscow, Bolshoy Trekhsvyatitelskiy side-street, 3

The paper describes the possibility of applying the model of situational control as a basis for creation adaptive movement control system of a walking robot. There is an example of creating system of situational control to select an optimal algorithm of movement of six-legged walking robot.

INTELLECTUAL TECHNOLOGIES FOR MANIPULATING TASKS

V.E. Pavlovsky (*vpavl@keldysh.ru*)
Keldysh Institute of Applied Mathematics of RAS, Moscow
E.A. Efimova (*yefi03@yandex.ru*)
A.P. Aliceychik (*atooxa@gmail.com*)
I.A. Orlov (*orlovbel@gmail.com*)
E.D. Elagina (*lizzaelagina@yandex.ru*)
Russian State University for the Humanities, Moscow.

In work the intellectual functions developed for robots manipulators are described. The universal planner intended for planning of operations of manipulators in complex environments is described, it is fulfilled on computer models of robots. The intelligent robot for playing logical board games is described, its logical program developed on the PROLOG is described. The created robots passed experimental working off which allowed to draw a conclusion on efficiency created program and hardware.

THE PROBLEM OF ANOMALY DETECTION IN TIME SERIES

Sergey G. Antipov (*antysergey@mail.ru*)
Vadim N. Vagin (*vagin@appmat.ru*)
Moscow Power Engineering Institute
111250 Krasnokazarmennaja str., 14, Moscow, Russia

In this paper the problem of anomaly detection in time series data is considered. The case when a learning set contains time series data concerning to some classes, is researched. The new algorithm TS-ADEEP-Multi for case when learning set contains some classes is suggested. The results of the program modeling are presented.

PROBLEM OF KNOWLEDGE GENERALIZATION BY PRESENCE OF NOISE IN CLASS ATTRIBUTES

Sergey G. Antipov (*antysergey@mail.ru*)
Ludmila A. Starostina (*StarostinaLA@mpei.ru*)
Marina V. Fomina (*m_fomina2000@mail.ru*)
Moscow Power Engineering Institute
111250 Moscow, Krasnokazarmennaja str., 14

The problem of knowledge generalization with the account of the necessity of processing the incomplete and inconsistent information stored in real databases is considered. It is offered to use methods of creation of production rules and decision trees in combination with different algorithms of processing of “noisy” values. Influence of noise on operation of algorithms of generalization on condition of entering of noise into the attribute defining a class of an example is researched.

INTELLIGENT WEB-SERVICE ENTERPRISE PERFORMANCE ANALYSIS

Tatiana V. Afanasieva (*tv.afanasjeva@gmail.com*)
Gleb Y. Guskov (*guskovgleb@gmail.com*)
Ulyanovsk state technical university, 32 SevernyVenetz str., 432027
Ulyanovsk, Russia
Irina G. Perfilieva (*irina.perfilieva@osu.cz*)
University of Ostrava, 30. dubna 22 701 03 Ostrava 1, Czech Republic
Anton A. Romanov (*romanov73@gmail.com*)
Irina A. Timina (*i.timina@ulstu.ru*)
Nadezhda G. Yarushkina (*jng@ulstu.ru*)
Ulyanovsk state technical university, 32 SevernyVenetz str.,
432027 Ulyanovsk, Russia

Choice of development strategy of a large enterprise depends on an assessment of its effectiveness. The article describes the software product that combines an expert system and a web service. Linguistic analysis and interpretation of non-linear dependencies of the source data (time series) is made using artificial intelligence and fuzzy logic methods, including the method of F-transform method and fuzzy trends.

SUPPORT VECTOR MACHINE DESIGN FOR SOLVING TEXT CATEGORIZATION PROBLEMS

Shakhnaz A. Akhmedova (*shahnaz@inbox.ru*)

Eugene S. Semekin (*eugeneseamenkin@yandex.ru*)

Siberian State Aerospace University
Krasnoyarsky rabochy avenue, 31, 660014,
Krasnoyarsk, Russia

Tatiana O. Gasanova (*tatiana.gasanova@ulm-uni.de*)

Wolfgang Minker (*wolfgang.minker@uni-ulm.de*)

Ulm University Albert-Einstein-Allee, 43, 89081 Ulm, Germany

New collective meta-heuristic and its modification for solving unconstrained and constrained optimization problems were developed. Proposed methods were used for tuning Support Vector Machine classifiers for solving text categorization problems. It was established that obtained classifiers is workable and effective for solving classification problems, they also demonstrated the performance higher than almost all known alternatives.

BIG DATA PROCESSING PROBLEMS AND SOLUTIONS

V.I. Gorodetsky (*gor@iias.spb.su*)

SPIIRAS, St. Petersburg

IT world is now entering the Big Data era. This phenomenon refers to the data explosion observed today. This term is widely used by scientists and practitioners. Big Data analysis provides for new opportunities in many respects. Analysis recent publications reveals that their most frequently themes are about new opportunities and challenges and very few, however, about solutions providing for predictive analytics going beyond the limits of OLAP models. The paper outlines the nature of challenges and some solutions proposed by the author to attack them.

THEORETICAL BASES OF THE METHODS OF CONSTRUCTION OF ASSOCIATION MEASURES IN INTELLIGENT DATA ANALYSIS

Ildar Z. Batyrshin (*batyr1@gmail.com*)

Centro de Investigacion en Computacion, Instituto Politecnico Nacional, Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, Delegación Gustavo A. Madero, C.P 07738, México D.F

The general methods of construction of functions satisfying axioms generalizing the properties of correlation coefficient are proposed. These methods give possibility to build non-statistical measures of association on sets with similarity measure and involution operation. The correlation coefficient is obtained as a special case of association measures related with Lukasiewicz t-conorm.

VKF-METHOD OF HYPOTHESES GENERATION: PROGRAMMING ISSUES

Dmitry V. Vinogradov (*vin@viniti.ru*)

All-russian institute for scientific and technical information RAS
20 Usievicha street, Moscow 125190, Russia

Programming and theoretical issues of VKF-method of hypotheses generation are represented. The approach combines induction, abduction and analogy procedures as JSM method does. Main novelties are Markov chain approach to induction step and the refinement procedure before abduction testing. The program was applied to SPECT dataset from UCI MLR. The VKF system uses the simplest strategy and demonstrates superiority over CLIP3 algorithm of the dataset's authors. It achieves 85.56% accuracy of prediction that exceeds 84% accuracy of CLIP3.

KNOWLEDGE DISCOVERY FOR THE PROBLEM OF COMPARING CURING STRATEGIES

Alexander I. Karachunskiy (*aikarat@mail.ru*)

Research and Clinical Center of Pediatric Hematology, Oncology and Immunology

Samora Machel st. 1, 117997 Moscow, Russia

Natalia V. Korepanova (*korepanova.natalia@gmail.com*)

Sergei O. Kuznetsov (*skuznetsov@hse.ru*)

National Research University Higher School of Economics
Bolshoy Tryokhsvyatitelsky Pereulok 3, Moscow, Russia

An approach to treatment-subgroup identification is proposed. This approach is based on matching algorithms and decision tree. An application to the data on children with acute lymphoblastic leukaemia is considered.

TRANSFORMING SPACES OF ATTRIBUTES WITH REGARD TO CONDUCT GENERALIZATIONS ERRORS

M.O. Korlyakova (*mkorlyakova@yandex.ru*)

A.Y. Pilipenko (*gamerjd@mail.ru*)

Kaluga Branch of Bauman Moscow State Technical University,
Kaluga

The results of analysis of the learning models in terms of homogeneity errors. Shown that the formation of the description of objects with low error variance can significantly improve the quality of the solution of the classification problem.

INCREASE OF JSM-METHOD PERFORMANCE FOR TEXT PROCESSING TASKS

Evgeniy V. Kotelnikov (*kotelnikov.ev@gmail.com*)

Vyatka State University of Humanities
610002, Kirov, Krasnoarmeyskaya, 26

The paper proposes the way of JSM-method performance increasing in text processing tasks based on a co-clustering of terms and documents using parallel computing. The co-clustering algorithm and parallelization scheme are described. The experimental results with text collections are shown.

CLASSIFICATION OF DIGRAPHS' SIMILARITY DETERMINING PROBLEMS

Victor V. Kokhov (*viktor-kokhov@ya.ru*)

Moscow Power Engineering Institute (Technical University)
111250, Moscow, Krasnokazarmennaya str., 14

Formalized statement of generalized problem of maximum common fragment determining and two digraphs similarity determining are defined. Classification of two digraphs' maximum common fragment determining problems is given. Classification of similarity determining problems is suggested. The universal method of the analyzed problem solving is declared.

INTELLIGENT SYSTEMS AND PROBLEMS OF SOCIAL RESEARCHES' SUPPORT

Victor K. Finn (*finn@viniti.ru*)

Maria A. Mikheyenkova (*mmikh@viniti.ru*)

All-Russian Institute for Scientific and Technical Information (VINITI), RAS;
125190, Usievitcha str., 20, Moscow, Russia

The problems of Intelligent Systems for researches in the social sciences are considered. The human-machine systems realize formalized heuristics simulating cognitive reasoning based on empirical data. The formation of the Base of facts is provided by thoughtful sociologist's work.

SOME RESULTS AND PROSPECT IN THE DEVELOPMENT OF CLASSIFICATION MODELS BASED GOOD CLASSIFICATION TESTS

Xenia A. Naidenova (*naidenovaxen@gmail.com*)

Military Medical Academy, Saint-Petersburg

A retrospective review of classification and pattern recognition development based on good classification tests is advanced. Some perspective directions of investigation are considered.

EVOLUTIONARY COVERING METHOD FOR COMPILING OF FACT BASE OF JSM-METHOD

Alexander Panov (*pan@isa.ru*)

Alexander Shvets (*shvets@isa.ru*)

Institute for Systems Analysis RAS

This paper presents a method to learn a compact fact base from a large dataset. Such a fact base is necessary to extract cause-and-effect relationships using JSM-method. A comparison of two methods for feature selection is performed. The first one applies AQ-based sequential covering, and the second one – evolutionary algorithm. Reducing feature space dimensionality is essential for JSM-method to perform within sensible time. The proposed methods were evaluated on psychological testing data.

**BAYESIAN BELIEF NETWORKS
FOR ESTIMATING OF RESPONDENTS' RISKY BEHAVIOR RATE**

Alena V. Suvorova (*suvalv@mail.ru*)

St. Petersburg Institute for Informatics and Automation
of RAS, St. Petersburg

Ruslan R. Khaibullin (*Ruslan.Khaibullin@lanit-tercom.com*)

St. Petersburg State University, St. Petersburg

Alexander L. Tulupyev (*alt@ias.spb.su*)

St. Petersburg Institute for Informatics and Automation
of RAS, St. Petersburg

We offered an approach to the risky behavior modelling on the base of data about behavior episodes. We described the developed software and the model based on Bayesian belief network and provided the results of applying the model to the statistical data.

**ALGORITHM FOR ANOMALY SEARCH IN THE PROCESSES
BASED ON FUZZY TRENDS IN TIME SERIES**

Tatiana V. Afanasieva (*tv.afanasjeva@gmail.com*)

D.V. Zavarzin (*dzavarzin91@gmail.com*)

Ulyanovsk state technical university, 32 SevernyVenetz str., 432027
Ulyanovsk, Russia

This paper proposes a new classification of possible anomalies in processes, the parameters of which can be represented by time series. Feature anomalies classification is to expand the traditional classification by fuzzy local trends. Formalized the conditions imposed by the detection of anomalies classes on time series and proposed a scheme of software algorithm.

TIME SERIES DYNAMICS FUZZY SCENARIO MODEL

M.M. Zernov (*zmmioml@yandex.ru*)

Smolensk Branch of the National Research University "Moscow Power
Engineering Institute"

1, Energeticheskij pr-d, Smolensk, Russian Federation, 214013

Time series dynamics fuzzy scenario model is offered in this paper. The model takes into consideration scenario realization uncertainty and allows predicting several phase of time series dynamics accordingly to identified scenario. Two methods of fuzzy scenario identification are developed.

FUZZY AND STOCHASTIC CONCLUSION IN INTELLECTUAL SYSTEMS OF PROCESSING OF PRIMARY INFORMATION

S.M. Kovalev (*ksm@rfnias.ru*)
Rostov State Transport University

In work the concept of intellectual pro-processing and cognitive measurements of information on the basis of use of schemes of an fuzzy and stochastic conclusion develops. The method is used in interests of creation of new classes of applied intellectual systems of dynamic type. The scheme of an fuzzy and stochastic conclusion for the indistinct dynamic systems allowing analytical representation in the form of functional dependences is offered.

FUZZY TRUTH VALUE INFERENCE METHOD FOR THE SYSTEMS WITH MULTIPLE FUZZY INPUTS

Dmitry Kutsenko (*dakutsenko@mail.ru*)
Vladimir Polakov (*p_v_m@mail.ru*)
Vasiliy Sinuk (*vgsinuk@mail.ru*)

Belgorod State Technological University named after V.G.Shoukhov
Russia, 308012, Belgorod, Kostyukov str., 46. Belgorod State Technological
University named after V.G.Shoukhov

The paper describes a fuzzy truth value inference method for fuzzy production models of systems with multiple fuzzy inputs. The method is compared with the original Zadeh's method and Mamdani's popular method. Proved computational efficiency of the proposed method.

RETE ALGORITHM MODIFICATION DEVELOPMENT FOR FUZZY EXPERT SYSTEMS

I.S. Mikhaylov (*fr82@mail.ru*)
National Research University Moscow Power Engineering Institute, Moscow
Zo Min Thaik (*zawgyi86@gmail.com*)
National Research University Moscow Power Engineering Institute, Moscow

In this paper basic concepts of fuzzy expert systems are considered. Developed modified algorithm Rete for the fuzzy rule base is proposed. This algorithm allows to formulate rules and conclusions using limited natural language and provides acceleration of the processing by same conditions single computation in different rules.

EXPERT SYSTEM ON THE BASIS OF FUZZY LOGIC AND NEURAL NETWORK TECHNOLOGIES

N.A. Polkovnikova (*npolkovnikova@tti.sfedu.ru*)

V.M. Kureichik (*kur@tgn.sfedu.ru*)

Engineering-technological academy of the Southern Federal University,
Taganrog

The paper shows development of fuzzy expert system model for fault identification of complex technical objects. Application of neural network technology can detect nonlinear dependence of input and output data, improve quality of object diagnostic process, which ultimately will reduce the number of accidents under operating conditions.

THIRD GENERATION KNOWLEDGE ACQUISITION SYSTEMS BASED ON COGNITIVE MEASUREMENTS

Maria N. Svyatkina (*maria.svyatkina@gmail.com*)

Valery B. Tarassov (*tarasov@rk9.bmstu.ru*)

Bauman Moscow State Technical University
5 b.1, 2nd Baumanskaya st., Moscow, 105005, Russia

Some models of non-classical measurements are analyzed in the context of designing 3rd generation knowledge acquisition systems. A new concept of cognitive measurements as a hierarchical information granulation process based on the principle of “Quantitative Measurement – Pragmatic Estimation” unity is proposed. A two-leveled architecture of cognitive measurements is considered. The logical-algebraic approach to data interpretation and analysis on the top level of cognitive measurements is developed.

ADJUSTMENT OF HYPOTHESES FORECAST FOR KNOWLEDGE EXTRACTION OF TIME SERIES

I.A. Timina (*i.timina@ulstu.ru*)

Ulyanovsk State Technical University, Ulyanovsk

In this paper the approach to the problem of the analysis and forecasting of time series included in the intellectual study of the time series model-based fuzzy time series models fuzzy trends associated with the adjustment of the hypotheses of the forecast and selected using clustering time series dominant fuzzy trends. Based on the results of the analysis of rows updated knowledge base time series of new rules.

FUZZY CLASSIFIER PARAMETER TUNING BASED ON CUCKOO SEARCH ALGORITHM

Ilya A. Hodashinsky (*hodashn@rambler.ru*)

Darya U. Minina (*pound_2007@mail.ru*)

Tomsk State University of Control Systems and RadioElectronics
40 Lenina, Tomsk 634050, Russia

In paper, the fuzzy rule-based classifier is considered. Parameters of the classifier are tuned through a cuckoo search algorithm. It is shown the performance of the proposed method on the four real world problems: balance, pima, bupa, iris data sets. The comparative analysis of the received results with analogues is given.

SEMANTIC TECHNOLOGIES OF QUESTION GENERATION IN LEARNING SYSTEMS

Madehur M. Ayupov (*madehur@mail.ru*)

Olga A. Nevzorova (*onevzoro@gmail.com*)

Nikolay A. Prokopjev (*nikolaskuji@gmail.com*)

Dzhavdet Sh. Suleymanov (*dvdtslt@gmail.com*)

Research Institute of Applied Semiotics of the Tatarstan Academy of Sciences,
Kazan Federal University, Kazan, Russia
Bauman Str. 20, 420111, Kazan, Russia

This paper describes the model of automatic question generation based upon the forms of all kinds of database queries. The ontological model of the “Informatics” subject area is represented in the tables of the relational database. Every test question is described at three levels: the ontological level, the lexical level and the level of implementation. The results of test question generation for the school course “Informatics” are also analyzed in this paper.

SUBJECTIVE CLASSIFICATION OF THE DOMAIN ONTOLOGIES

Tatiana Gavrilova (*gavrilova@gsom.pu.ru*)

Graduate School of Management St. Petersburg University
199004 St. Petersburg, Russia Volkhovsky per., 1-3

The paper presents the main methodological problems of ontology design and development. The new classification based on different criteria of ontology assessment is proposed. The stress is put on possible mistakes and errors in practical domain ontologies design.

A METHOD FOR DESIGN OF RULE-BASED EXPERT SYSTEMS ON THE BASIS OF MODEL DRIVEN APPROACH

Maksim A. Grishchenko (*makcmg@icc.ru*)

Olga A. Nikolaychuk (*nikoly@icc.ru*)

Alexander Yu. Yurin (*iskander@icc.ru*)

Institute for System Dynamics and Control Theory, Russian Academy of Sciences, Siberian Branch, 664033, Lermontov st., 134, Irkutsk, Russia

The paper describes a modification of a model-driven approach, in particular, the MDA (Model Driven Architecture) methodology. The proposed modification provides to develop rule-based expert systems. The statement of the problem is made. The main stages and templates of rules for transformation of models are described.

TOWARDS SEMANTIC GEO-INTEROPERABILITY IN GEOSPATIAL SEMANTIC WEB

Sergey K. Dulin (*s.dulin@ccas.ru*)

Dorodnicyn Computing Centre of the Russian Academy of Sciences
119333, Moscow, Vavilova str., 40

Natalia G. Dulina (*ngdulina@mail.ru*)

Dorodnicyn Computing Centre of the Russian Academy of Sciences
119333, Moscow, Vavilova str., 40

Olga S. Kozhunova (*kozhunovka@mail.ru*)

The Institute of Informatics Problems of the Russian Academy of Sciences
119333, Moscow, Vavilova str., 44, bldg.2

The task of semantic geo-interoperability is to ensure coherent interaction of experts to solve problems that require sharing geo resources, provided an adequate understanding of the semantics embedded in the spatial ontology and/or geo thesauruses and classifiers. This paper is devoted to the discussion of the notion of semantic interoperability and aspects of its implementation.

TECHNOLOGY ROADMAPS. FORMAL MODELS OF BEST PRACTICES

I.V. Efimenko (*iefimenko@hse.ru*)

Higher School of Economics, Moscow

V.F. Khoroshevsky (*khos@ccas.ru*)

Dorodnitsyn Computing Centre RAS, Moscow

The paper represents a formal model of technological roadmapping (TRM). The research was carried out based on the analysis of best practices in the domain. It integrates the experience of a wide range of TRM projects. The proposed model provides basis for developing tools for TRM automation, both for producing stage and usage. It can be also applied to creating standardized roadmaps. A system of ontologies specifies key concepts of TRM and forms the static component of the model. Roadmap architecture (formats, main tasks being solved, etc.), frequently used types of layers and elements are considered. The dynamic component is implemented within a process model integrating a set of high-level algorithm schemes. Developing a guide which formalizes recommendations for TRM builders, and a range of knowledge-intensive tasks such as knowledge extraction in TRM domain as well as automated quantification of qualitative information used in foresight are considered for further research.

THE SUPPORT OF KNOWLEDGE BASES CONTROL FOR AUTOMATION OF INTELLIGENT PROFESSIONAL ACTIVITY

A.S. Kleshev (*kleshev@iacp.dvo.ru*)

E.A. Shalfeeva (*shalft@iacp.dvo.ru*)

the Institute for Automation & Control Processes of the FEB RAS

This article is devoted to organization of control process of quality of knowledge bases used for support of decision-making in daily intelligent activity. The outlines of automation of control process of knowledge bases quality and the main objectives of program systems for knowledge bases quality control are offered.

ALGORITHMS OF OPTIMAL SCHEDULE INFORMATION IN THE FORM OF KNOWLEDGE FOR NETWORKS WITH THE TREE TOPOLOGY

Alexander A. Kobayakov (*cri-granit@peterlink.ru*)

Kirill V. Lapshin (*kir_i_k@mail.ru*)

Yuriy A. Yamshchikov (*gcabman@yandex.ru*)

FSPC JSC "Concern "Granit-Electron"

Russia, 191014, Saint-Petersburg, Gospitalnaia, 3

This article discusses the tree sensor networks of the intellectual integrated multifunctional information management systems. For each of the subsets of a tree network is calculated lower limit of the length of time gathering information and proposed the algorithm of optimal routing of information in the form of knowledge that can be used for any tree sensor network.

ONTOLOGY-BASED INTEGRATION FOR RELATIONAL DATA

Alexander O. Kolesov (*alexander.kolesov90@gmail.com*)

Alexey M. Namestnikov (*nam@ulstu.ru*)

Ulyanovsk State Technical University, Ulyanovsk

This article describes principles of work, architecture and implementation of Ontology-based Integration system for relational data. It also shows how to use the system on the basis of experimental data.

PROBLEMS AND METHODS OF TEMPORAL DIGRAPHS' SIMILARITY ANALYSIS

Victor V. Kokhov (*viktor-kokhov@ya.ru*)

Moscow Power Engineering Institute (Technical University)

111250, Moscow, Krasnokazarmennaya str., 14

Two classes of determining the similarity problem of two temporal digraphs for the purpose of temporal digraphs' structural properties dynamics change research are defined. Two approaches in development of the problem solution methods are suggested. Comparative analysis of methods from two suggested approaches is given.

THE SOFTWARE FOR LINGUISTIC-DIRECTED COMPLETION OF ONTOLOGY

Serge S. Kurbatov (*curbatow.serg@yandex.ru*)

Alex P. Lobzin (*lobzin@rambler.ru*)

Gennady K. Khakhalin (*gkhakhalin@yandex.ru*)

Research center for electronic computer technology

NICEVT, Warsawskoye highway, 125, 117178,

Moscow, Russia

The logic of software creation oriented on ontology completion of means of linguistic translation is described. The interface with ontology providing correspondence of NL descriptions and formal constructions is offered. The mechanism of assumption advancement on generalization of the entered information is developed. The breadboard model embodying developed logic in programs is realized.

**FEATURES OF THE INTEGRATION OF THE INFERENCE
ENGINE IN AN ONTOLOGICAL MODEL OF KNOWLEDGE
REPRESENTATION WITH SWRL.**

Vadim S. Moshkin (*postforvadim@ya.ru*)
Nadezhda G. Yarushkina (*jng@ulstu.ru*)
Ulyanovsk state technical university
432027, Ulyanovsk, Severny Venec st., 32

This article provides a semantic approach summarizing time series as an example of indicators of local area network through the use of ontology problem area. Solved the problem of integrating the various approaches provide expert knowledge through the integration of knowledge in the ontological nature of a production model using SWRL-rules.

**INTELLIGENT DECISION SUPPORT SYSTEM FOR
SUBSTANTIATION OF STRUCTURAL MATERIALS**

Alexander F. Berman (*berman@icc.ru*),
Galina S. Maltugueva (*gama@icc.ru*),
Alexander Yu. Yurin (*iskander@icc.ru*)
Institute for System Dynamics and Control Theory, Russian Academy of
Sciences, Siberian Branch,
664033, Lermontov st., 134, Irkutsk, Russia

The paper describes an intelligent decision support system (IDSS) for substantiation of structural materials. The algorithms implemented in the IDSS are based on the combined use of case-based reasoning and methods of multicriteria decision-making. The model of case and the example of application of IDSS are presented.

**THE PROTOTYPE OF GENETIC ALGORITHM IMPLEMENTATION
FOR DOCUMENT RETRIEVAL**

Vladimir K. Ivanov (*mtivk@mail.ru*)
Boris V. Palyukh (*pboris@tstu.tver.ru*)
Pavel .I. Meskin (*pavel.meskin@gmail.com*)
Tver State Technical University
Russian Federation, Tver, Afanasy Nikitin st. 22

This article describes the prototype of genetic algorithm implementation for identifying and selecting the most relevant search results obtained during document

retrieval by search engines. The fitness function, basic steps and parameters of the algorithm are considered. Main components of software are described. Also some results of study of algorithm behavior are presented.

BUILDING OF KNOWLEDGE PORTAL FOR INFORMATION AND ANALYTICAL SUPPORT OF DSS DEVELOPERS

G.B. Zagorulko (*galr@iis.nsk.su*)

A.P. Ershov Institute of Informatics Systems, Novosibirsk

V.G. Itygilov (*iti.vadim@gmail.com*)

Novosibirsk State University, Novosibirsk

The paper describes the development of a knowledge portal which systematizes available information about decision support methods on the base of ontology and provides tools for viewing, analysis and practical use of this information.

APPROACH TO AUTOMATIZATION OF INFORMATION COLLECTION FOR THEMATIC INTELLIGENT SCIENTIFIC INTERNET RESOURCES

Yu.A. Zagorulko (*zagor@iis.nsk.su*)

A.P. Ershov Institute of Informatics Systems, Novosibirsk

O.I. Borovikova (*olesya@iis.nsk.su*)

A.P. Ershov Institute of Informatics Systems, Novosibirsk

E.A. Sidorova (*lena@iis.nsk.su*)

A.P. Ershov Institute of Informatics Systems, Novosibirsk

I.R. Ahmadeeva (*ah.irishka@gmail.com*)

Novosibirsk State University, Novosibirsk

The paper presents an ontology-based approach to automatization of information collection for thematic intelligent scientific internet resources providing the systematization and integration of scientific knowledge, information resources, related to certain subject domain, and methods of intelligent processing of data contained in them as well as the content-based access to them.

INTELLIGENT SYSTEM for CYBER SECURITY MONITORING AND INCIDENT MANAGEMENT

Igor V. Kotenko (*ivkote@comsec.spb.ru*)

Igor B. Saenko (*ibsaen@comsec.spb.ru*)

Saint-Petersburg Institute for Information and Automation of Russian Academy
of Sciences

The paper considers methods and architecture of intellectual system for cyber security monitoring and incident management, which extends the capabilities of traditional systems of this type. A formal description of the methods used in the early stages of monitoring and for decision-making is outlined. An example of their realization, the implemented system architecture and its testing areas are considered.

ONTOLOGICAL SYSTEM FOR THE INTELLECTUAL ANALYSIS OF TECHNICAL DOCUMENTATION

Ruslan A. Subkhangulov (*subkhangulov-ruslan@yandex.ru*)

Aleksey A. Filippov (*al.filippov@ulstu.ru*)

Ulyanovsk State Technical University
32, Severny Venetz str.,432027 Ulyanovsk, Russia

In article the description of program system for structuring and information retrieval of technical text documents is provided. The method based on a current status of domain ontology. The structure of ontology is shown, basic functions of system are listed. The results of the experiments proving efficiency of developed program system are given.

TOWARDS AUTOMATIC ASSIGNMENT GENERATION FOR A CROWDSOURCED THESAURUS

D.A. Ustalov (*dau@imm.uran.ru*)

IMM UB RAS, Ekaterinburg

620990 S. Kovalevskaya, 16, Ekaterinburg, Sverdlovskaya obl., Russia

Automatic generation of user assignments is an important problem of collectively created information resources. They involve a large amount of volunteers to participate in annotation. In this paper such a problem was analyzed and a method of solution has been proposed using an open electronic thesaurus for Russian project as the example.

IMAGE ANALYSIS OBTAINED FROM A CAMERA ONBOARD THE UNMANNED AIRCRAFT FOR ITS NAVIGATION

Hussein A. Abdulkadhim (*h_a_meer@yahoo.com*)

Ulyanovsk State Technical University, Ulyanovsk.

This article presents some results of research of the possibility of applying different algorithms for estimating of spatial displacements of the images obtained from the onboard camera autonomous UAV, for solving problems of an autonomous vehicle navigation.

DEVELOPMENT OF ALGORITHMS AND CONTROL SOFTWARE OF UNMANNED AERIAL VEHICLES

S.V. Voronov (*s.voronov@ulstu.ru*)
V.E. Dement'ev (*vitawed@mail.ru*)
S.S. Loginov (*ss7loginov@gmail.com*)
Ulyanovsk State Technical University

This paper considers the solution of the control problem of unmanned aerial vehicle based on the use of interconnecting inertial navigation systems and data from the on-board video camera. It also describes the control software.

DEVELOPMENT OF ALGORITHMS NAVIGATION AUTONOMOUS AIRCRAFT BASED ON THE AGGREGATION OF DIFFERENT NAVIGATION DATA

Vitaliy E. Dementyev (*vitawed@mail.ru*),
Dmitriy S. Kondratyev (*kondratev.dmitriy@gmail.com*)
ULSTU, Ulyanovsk (*rector@ulstu.ru*)

The paper describes the use of search algorithms for objects in the image obtained with the infrared range finder, as well as the aggregation of this data with on-board inertial systems in order to determine its own location.

UNMANNED VEHICLE – THE ACTUAL STATE AND PERSPECTIVE

Dmitriy A. Dobrynin (*dobr@viniti.ru*)
VINITI RAS, Moscow

The paper considers an application of AI for building unmanned vehicle. Describes the actual state and perspective of development for intellectual transport robots.

VISUAL NAVIGATION OF UNMANNED AERIAL VEHICLES IN AN UNCERTAIN AND VARIABLE CONDITIONS OF OBSERVATION

N. Kim (*nkim2011@list.ru*)
Moscow Aviation Institute, Moscow
N. Bodunkov (*boduncov63@yandex.ru*)
Moscow Aviation Institute, Moscow

This paper examines problems associated with the organization of the autonomous visual navigation of the unmanned aerial vehicles in variable and uncertain conditions of observation. A method and an adaptive algorithm to recognize landmarks are proposed. Experimental results that demonstrate effectiveness of the proposed method are demonstrated

COORDINATED CONTROL OF AUTONOMOUS GROUPS OF UAVs

N. Kim (*nkim2011@list.ru*)
Moscow Aviation Institute, Moscow
I. Krylov (*Krylov_I@mail.ru*)
Moscow Aviation Institute, Moscow

The paper describes an algorithm for formation of coordinated control for unmanned aerial vehicles for group searching ground objects

INTELLIGENT INFORMATION PROCESSING IN COMPUTER VISION SYSTEM

M.O. Korlyakova (*mkorlyakova@yandex.ru*)
P.K. Novikov (*bobersuperzy_b@mail.ru*)
E.Y. Prokopov (*evgen-prokopov@rambler.ru*)
Kaluga Branch of Bauman Moscow State Technical University, Kaluga

This work describes approach to intelligent information processing in computer vision system. The main process in image recognition and stereo reconstruction was present. The samples of same stage of image processing in neural network was described.

THE CONTROL AND INFORMATION SUPPORT PROBLEMS OF GROUPS OF DRONES USAGE

A.A. Lipatov (*my-post-box@hotmail.ru*)
D.A. Milyakov (*from_fn@mail.ru*)
JSC Vega Radio Engineering Corporation, Moscow

This paper contain the analysis of a number of control and information support problems that arise from the use of groups of drones. The methods of solving problems of optimal centralised control of drones group trajectory and spatial order, and identification of such groups with the use of uncertain quantitative and qualitative information are considered.

RESEARCH OF MOVEMENT OF JUMPING ROBOT EQUIPPED WITH WINGS WITH VARIABLE GEOMETRY

Oksana Gennadievna Loktionova (*teormeh@inbox.ru*)
Lyudmila Yurievna Vorochaeva (*mila180888@yandex.ru*)
Alexander Valerievich Vorochaev (*sasha-vorochaev@yandex.ru*)
South-West State University, Kursk, 50 let Oktyabry, 94

In the article considers jumping-flying robot, which wings enable the planning of the device during flight stage, developed mathematical model of movement of the device, particular attention is paid to the interaction of the wings with the environment.

MODEL, BASIC TRAJECTORIES, NEUROCONTROLLER FOR MULTI ROTOR ROBOT

V.E. Pavlovsky (*vlpavl@mail.ru*)
M.V. Keldysh Institute of Applied Mechanics of RAS, Moscow
A.V. Savitskij (*as@ttorr.ru*)
Lomonosov Moscow State University, Moscow

This article describes the theoretical model of multirotor robot - quadcopter with the basic mechanical effects. It was studied the basic trajectories, including takeoff, landing and flight in a straight line. On the strength of the basic numerical simulations it was constructed the neurocontroller for the quadcopter.

SOFTWARE FOR SIMULATION OF AUTONOMOUS UAVs IN GROUP OPERATIONS

Nikolay V. Kokhtenko (*nvkokhtenko@gmail.com*)
ZAO NII TS "Sinvent"
Moskovsky pr. 212, Saint-Petersburg, 196066, Russia
Sergey V. Lebedev (*lebedev.sv.etu@gmail.com*)
Saint-Petersburg State Electrotechnical University "LETI"
Prof. Popova st. 5, Saint-Petersburg, 197376, Russia
Michael G. Panteleyev (*MPanteleyev@gmail.com*)
Saint-Petersburg State Electrotechnical University "LETI",
OAO NIC SPb ETU,
Prof. Popova st. 5, Saint-Petersburg, 197376, Russia

The problems of creating a software for simulation of UAVs with a high degree of autonomy in group operations are discussed. The software is intended for experimental research and real-time testing of various models of UAV's group

behavior under team counteractions. Requirements to the software are formulated. An architecture and main features of the functioning of the software are considered.

SOFTWARE SYSTEMS FOR UAV FLIGHT SIMULATION AND ITS AUTONOMOUS POSITIONING USING VISION METHODS AND ALGORITHMS

D.N. Stepanov (*mitek1989@mail.ru*)

I.P. Tishchenko (*billy@billy.botik.ru*)

Program System Institute named A.K. Aylamazyan of RAS,
Pereslavl-Zalessky

The paper describes a software system designed to simulate the flight of unmanned aerial vehicle with the help of tools, methods and algorithms of technical vision. The general scheme of the generation and processing of data, as well as a brief description of the modules of a software system designed to solving the individual subtasks, are described.

THE TASK OF UNMANNED AERIAL VEHICLE FLYING AROUND TERRITORY SUBJECT TO WIND LOADS

M.V. Khachumov (*khmike@inbox.ru*)

Institute for Systems Analysis of Russian Academy of Sciences, 117312,
Moscow, pr. 60-letiya Oktyabrya, 9

In the article the task of the unmanned aerial vehicle flying around territory defined by reference points, subject to wind loads and control restrictions is solved. The problem is divided into the task of flying trajectory planning and the task of following the chosen route. In accordance with the principle of separation of motion the flat case is considered. The solution of the task of following is based on production rules.

DISTRIBUTED INTELLIGENT CONTROL SYSTEM FOR SQUADS OF UAVS FUNCTIONING IN GPS-DENIED ENVIRONMENT

K.S. Yakovlev (*yakovlev@isa.ru*)

Institute for Systems Analysis of Russian Academy of Sciences, Moscow

V.V. Khitkov (*vskhitkov@gmail.com*)

M.I. Loginov (*lunarstrainut@gmail.com*)

A.V. Petrov (*gmdidro@gmail.com*)

Soloviev Rybinsk State Aviation Technical University, Rybinsk

Distributed intelligent system controlling a squad of quadrotor UAVs (AR. Drone 2.0) functioning in GPS-denied environment and solving a variation of visual

SLAM task is presented in the paper. Control system architecture is based on 4D/RCS framework and is implemented within ROS paradigm. Given task is solved with the EKF-SLAM algorithm.

RESEARCH OF INFLUENCE OF WINGS'S LENGTHS OF THE ORNITOPTER ON HIS KINEMATICS FEATURES

Sergey Fedorovich Jatsun (*teormeh@inbox.ru*)

South-West State University, Kursk

Sergey Venegditovich Efimov (*scione@yandex.ru*)

Voronezh Institute of Russian Ministry for Emergency Situations

Grigorii Sergeevich Naumov (*gregory-XXI@yandex.ru*)

South-West State University, Kursk

In the article considers the ornitopter – flying robot with waving wing, imitating moving the bird during flight. The mathematical model of the object is building, the result in type graph are presented. Influence of the lengths section wing on average velocity of the robot is explored.

MODELLING OF UNMANNED QUADROTOR AIRCRAFT FLIGHT

S.F. Jatsun (*teormeh@inbox.ru*)

N.I. Popov, O.V. Emelianova, A.I. Savin

Southwest State University, Kursk, Russia

The paper considers questions of quadcopter movement mathematical modeling, taking into account mass-dimensional properties of four electrical drives with mechanical reducer. The design diagram and the differential equations based on the general theorems of dynamics describe the interconnected electromagnetic and mechanical processes in electromechanical system of quadcopter propeller drives. The offered nonlinear differential equations are solved with the kinematic ratios which express angular speed projections of a body on the related coordinate system axes via time derivatives of angles: roll, pitch and yaw.