

Міністерство освіти і науки України
Національна академія наук України
Національний авіаційний університет
Інститут комп'ютерних технологій



МІЖНАРОДНА НАУКОВО-ТЕХНІЧНА
КОНФЕРЕНЦІЯ
«ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ
ЛІНГВІСТИЧНОГО АНАЛІЗУ»

21-22 жовтня 2009 року

Тези доповідей

Київ 2009

УДК 004.01(082)

Міжнародна наукова-технічна конференція «Інтелектуальні технології лінгвістичного аналізу»: Тези доповідей. – К.: Вид-во Нац. авіац. ун-ту «НАУ-друк», 2009. – 76 с.

Збірник містить тези доповідей, які були представлені на конференції «Інтелектуальні технології лінгвістичного аналізу».

В доповідях розглянуті наукові та методичні питання інтелектуальних технологій: методологія інтелектуальних мовно-інформаційних систем, комп'ютерна технологія порівняльного аналізу електронних текстів, технології інформаційного пошуку, методи системного моделювання. Для фахівців з комп'ютерної лінгвістики.

Редакційна колегія:

Литвиненко О.Є. – д.т.н., професор, декан факультету комп'ютерних систем Інституту комп'ютерних технологій НАУ

Литвинов В.В. – д.т.н., професор, начальник відділу Інституту проблем математичних машин і систем НАН України

Денисюк В.П. – д.ф.-м.н., професор, завідувач кафедри вищої та обчислювальної математики факультету комп'ютерних систем Інституту комп'ютерних технологій НАУ

*Затверджено до друку вченою радою Інституту комп'ютерних технологій Національного авіаційного університету
(протокол № 13 від 29.10.2009 р.)*

Свідоцтво про державну реєстрацію ВОО №904384 від 13.06.05

© Національний авіаційний університет, 2009

ЗМІСТ

Д.В. Ланде, В.Ю. Зубок

ЛІНГВІСТИЧНИЙ ПІДХІД ДО ВИЯВЛЕННЯ ІНФОРМАЦІЙНИХ АТАК
З МЕРЕЖІ ІНТЕРНЕТ 9

М.Ю. Сич

СТРУКТУРА ТА МЕХАНІЗМИ РОБОТИ ІНТЕЛЕКТУАЛЬНОЇ
ІНФОРМАЦІЙНО-ДОВІДКОВОЇ СИСТЕМИ «ВІКТОРІЯ»..... 10

L.N. Badorina

MODEL COUSTUCTION OF FORMAL CORRELATION 11

Ю.О. Янов

МЕТОД ПОСЛІДОВНОЇ АДАПТАЦІЇ ЗАГАЛЬНИХ ПРАВИЛ
ВИЯВЛЕННЯ ТА ПРЕДСТАВЛЕННЯ ІНФОРМАЦІЇ ДЛЯ
ІНФОРМАЦІЙНО-АНАЛІТИЧНИХ СИСТЕМ 12

О.В. Ругайн

ДИДАКТИЧНІ АСПЕКТИ КОМП'ЮТЕРНОЇ ОБРОБКИ ПРИРОДНОЇ
МОВИ 13

А.І. Вавіленкова

СИСТЕМА АВТОМАТИЗОВАНОГО ФОРМУВАННЯ ЛОГІКО-
ЛІНГВІСТИЧНИХ МОДЕЛЕЙ ТЕКСТОВОЇ ІНФОРМАЦІЇ 14

А.С. Савостеня

АКТУАЛЬНІ ПИТАННЯ МОРФОЛОГІЧНОГО СИНТЕЗУ ТЕКСТІВ
УКРАЇНСЬКОЇ МОВИ 15

А.С. Шевченко

СТРАТЕГІЇ ІНФОРМАЦІЙНОГО ПОШУКУ 16

О.О. Добріна

МЕТОДИ ЗАХИСТУ АВТОРСЬКОГО ПРАВА НА ЗВУКОВІ ФАЙЛИ. 17

К.Ю. Бабий

СПОСОБ ЛОКАЛИЗАЦИИ ФУНКЦИОНАЛЬНОСТИ В ОТДЕЛЬНЫЕ
МОДУЛИ..... 18

А.А. Остапов, А.О. Соболев

СОЗДАНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ТЕСТИРОВАНИЯ
КАЧЕСТВА ФУНКЦИОНИРОВАНИЯ СЕТЕЙ 19

М.П. Чайка	
ОБРАХУНОК ЕЛЕМЕНТАРНИХ ФУНКЦІЙ	20
І.А. Клименко, С.М. Біляєв	
РЕАЛІЗАЦІЯ СХЕМИ УПРАВЛІННЯ СТАНАМИ ТА ЗСУВАМИ НА ПЛІС.....	21
О.С. Папук, І.О. Ніколаєнко, А.М. Алексійчук	
ІСТОРІЯ РОЗВИТКУ ТА НАПРЯМИ КОМП'ЮТЕРНОЇ ЛІНГВІСТИКИ.	22
Ю.М. Козир	
ПОШУК ІНФОРМАЦІЇ ЗАДАНОЇ ТЕМАТИКИ	23
О.О. Білик	
АЛГОРИТМ НЕЧІТКОГО ПОШУКУ ІНФОРМАЦІЇ.....	24
И.О. Николаенко, И.В. Ефименко	
ИНФОРМАЦИОННЫЕ И СОДЕРЖАТЕЛЬНЫЕ ОСНОВЫ ДЕЛОВОГО ПЕРЕВОДА В ОБЛАСТИ МУЛЬТИМЕДИА	25
О.Є. Литвиненко	
ПІДСИСТЕМА ПОРІВНЯЛЬНОГО АНАЛІЗУ ЕЛЕКТРОННИХ ТЕКСТІВ ЯК ЗАСІБ ПІДВИЩЕННЯ ЯКОСТІ ПІДГОТОВКИ ФАХІВЦІВ	26
Т.В. Шепелева, Г.С. Краліна	
ДОСЛІДЖЕННЯ І РОЗРОБКА ЕФЕКТИВНИХ АЛГОРИТМІВ ТОКЕНІЗАЦІЇ.....	27
М.В. Лупандін, І.А. Безкровна, І.М. Юр'єва	
РОЗРОБКА АЛГОРИТМІВ СЕМАНТИЧНОЇ СТРУКТУРИЗАЦІЇ ДАНИХ.	28
О.М. Дишлок, О.П. Нечипорук	
РОЗРОБКА МЕТОДІВ ТА АЛГОРИТМІВ ІНДЕКСАЦІЇ ЕЛЕКТРОННИХ ТЕКСТІВ	29
В.И. Надточий, Е.Д. Надточий,	
АЛГОРИТМИЗАЦИЯ ПРОЦЕССА АНАЛИЗА ЯЗЫКОВЫХ ЯВЛЕНИЙ.	30
Д.Ю. Катюха	
ОБ'ЄКТНИЙ ПІДХІД ДО КОМП'ЮТЕРНОГО АНАЛІЗУ ТЕКСТІВ ЗА ЗМІСТОМ.....	31
А.А. Беляков	
ИСПОЛЬЗОВАНИЕ КЛЮЧЕВЫХ СЛОВ И ВНЕШНИХ ТЕМАТИЧЕСКИХ ССЫЛОК ДЛЯ ПОИСКОВОЙ ОПТИМИЗАЦИИ САЙТОВ.....	32

Є.Б. Артамонов

КОНЦЕПЦІЇ СТВОРЕННЯ ЕЛЕКТРОННИХ ПІДРУЧНИКІВ З
МОЖЛИВІСТЮ АВТОМАТИЧНОГО НАЛАШТУВАННЯ ГЛИБИНИ
ПРЕДСТАВЛЕННЯ ІНФОРМАЦІЇ..... 33

Т.Г. Куцар

ВИЛУЧЕННЯ ФАКТОГРАФІЧНОЇ ІНФОРМАЦІЇ В ТЕКСТОВИХ
ДОКУМЕНТАХ 34

С.М. Денисенко

АНАЛІЗ ПАРАМЕТРІВ МОВИ У КОМП'ЮТЕРНИХ СИСТЕМАХ
ОБРОБКИ СИГНАЛІВ..... 35

Д.В. Ланде, В.В.Жигало

МЕТАПОИСКОВАЯ СИСТЕМА PDF SCIENCE SEARCH 36

В.В. Ковальчук

ЗАСОБИ ПІДВИЩЕННЯ РЕЛЕВАНТНОСТІ ПОШУКУ В
ІНФОРМАЦІЙНИХ СИСТЕМАХ НАУКОВИХ БІБЛІОТЕК 37

Е.Ю. Романец, А.В. Лищинский

ТЕОРИЯ И ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО ПОИСКА..... 38

Т.О. Слухай

МЕТА І ЗАВДАННЯ ДОСЛІДЖЕННЯ ТЕХНОЛОГІЙ
ІНФОРМАЦІЙНОГО ПОШУКУ 39

Д.А. Бурко

ОПТИМІЗАЦІЯ МОДУЛЯ ПОШУКУ ДАНИХ В ІНФОРМАЦІЙНИХ
СИСТЕМАХ 40

К.А. Мацуева

ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО ПОИСКА (ТЕМАТИЧЕСКИЙ
ПОИСК)..... 41

Н.П. Кадет, М.В. Сорока

ВИКОРИСТАННЯ СЕМАНТИЧНОГО ПОШУКУ В ІНФОРМАЦІЙНИХ
СИСТЕМАХ 42

А.Ю. Михайлюк, О.В. Пилипчук, М.В. Сніжко, В.П.Тарасенко

СПОСІБ СТРУКТУРНО-АЛГОРИТМІЧНОЇ ОРГАНІЗАЦІЇ
ІНТЕЛЕКТУАЛЬНОГО РЕДАКТОРА ЗАПИТУ НА
КВАЗІСЕМАНТИЧНИЙ ПОШУК..... 43

М.А. Мелешко, С.М.Денисенко, Ю.В. Помазков МОДЕЛЬ СУБ'ЄКТНОГО ОТОЧЕННЯ ВІДКРИТОЇ АРХІВНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ	44
О.В. Козлова МЕТОДИ ІНФОРМАЦІЙНОГО ПОШУКУ В ЕЛЕКТРОННИХ БІБЛІОТЕКАХ	45
Н. Л. Соколова СЕМАНТИЧНИЙ МЕХАНІЗМ В ТЕХНОЛОГІЯХ ІНФОРМАЦІЙНОГО ПОШУКУ	46
G.S. Kralina, E.V. Turota THE INFORMATION RETRIEVAL TECHNOLOGY	47
Д.С. Газізов, О.І. Курченко ПОШУК ІНФОРМАЦІЇ, РЕЛЕВАНТНОЇ ПОТРЕБАМ КОРИСТУВАЧА	48
Л.И. Гришина, Б.Н. Нестеренко, И.А. Шахматов ПОСТАНОВКА ЗАДАЧИ ДИСПЕТЧЕРСКОГО УПРАВЛЕНИЯ ПАРКОМ МАШИННО-ТЕХНОЛОГИЧЕСКИХ СИСТЕМ	49
М.О. Ментю АКТУАЛЬНІСТЬ ПРОБЛЕМИ ПОШУКУ ІНФОРМАЦІЇ	50
Е.А. Зубарева МЕТОД ПЕРЕДАЧИ МУЛЬТИМЕДІЙНОГО ТРАФИКА В БЕСПРОВОДНЫХ МУЛЬТИСЕРВИСНЫХ СЕТЯХ	51
О.Є. Литвиненко, А.В. Полухін, А.Б. Чапкіс ПРИНЦИПИ ПОБУДОВИ ТИПОВОЇ КОМП'ЮТЕРИЗОВАНОЇ СИСТЕМИ ПОТОЧНОГО ТА ПІДСУМКОВОГО КОНТРОЛЮ ЗНАНЬ СТУДЕНТІВ	52
І.А. Жуков, І.А. Клименко, В.О. Саченко МЕТОД РЕПЛІКАЦІЇ ДАНИХ В МОБІЛЬНИХ GRID	53
В.Є. Снитюк ЕВОЛЮЦІЙНІ ТЕХНОЛОГІЇ ПРИЙНЯТТЯ РІШЕНЬ В УМОВАХ НЕПОВНОЇ АПРІОРНОЇ ІНФОРМАЦІЇ	54
И.И. Горбань ФИЗИКО-МАТЕМАТИЧЕСКАЯ ТЕОРИЯ ГИПЕРСЛУЧАЙНЫХ ЯВЛЕНИЙ	55

В.В. Ткаченко, А.В. Левчук, Д.С. Пономарчук ВИРШЕННЯ ЗАДАЧ УПРАВЛІННЯ З ЗАСТОСУВАННЯМ КЛАСТЕРНИХ ОБЧИСЛЕНЬ	56
Т.В. Шепелева, А.О. Медведчук АРХІТЕКТУРА ТИПОВОЇ КОМП'ЮТЕРИЗОВАНОЇ СИСТЕМИ ПОТОЧНОГО ТА ПІДСУМКОВОГО КОНТРОЛЮ ЗНАНЬ СТУДЕНТІВ ВИЩИХ НАВЧАЛЬНИХ ЗАКЛАДІВ.....	57
В.В. Жабина ОРГАНИЗАЦІЯ ПАРАЛЛЕЛЬНИХ ПОТОКОВ КОМАНД В СИСТЕМАХ, УПРАВЛЯЕМЫХ ПОТОКОМ ДАННЫХ.....	58
В.И Жабин, В.В. Макаров ТАБЛИЧНО-АЛГОРИТМИЧЕСКИЙ МЕТОД ПОЛИНОМИАЛЬНОЙ АППРОКСИМАЦИИ ФУНКЦИЙ	59
О.М Глазок. ЗАСТОСУВАННЯ ДРУГОГО МЕТОДУ ЛЯПУНОВА ДЛЯ РОЗВ'ЯЗАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ ВЕЛИКОЇ РОЗМІРНОСТІ.....	60
Є.В. Шевцова СПОСІБ ФУНКЦІОНУВАННЯ СИСТЕМИ МОНИТОРИНГУ ТРАФІКУ БЕЗДРОТОВИХ МУЛЬТИСЕРВІСНИХ МЕРЕЖ.....	61
Д.В. Лобурь МОНИТОРИНГ ХАРАКТЕРИСТИК СУПУТНИКОВИХ НАВІГАЦІЙНИХ СИСТЕМ	62
С.С. Донченко ОБЕСПЕЧЕНИЕ ЗАДАНЫХ ПАРАМЕТРОВ КАЧЕСТВА ОБСЛУЖИВАНИЯ ДЛЯ ТРАФИКА РЕАЛЬНОГО ВРЕМЕНИ	63
О.М. Дишлюк МАТЕМАТИЧНА МОДЕЛЬ РОЗВ'ЯЗАННЯ ЗАДАЧІ АНАЛІЗУ РОБОТИ СОТОВИХ СИСТЕМ ЗВ'ЯЗКУ CDMA-ТЕХНОЛОГІЇ	64
И.И. Давиденко СПОСОБ ОРГАНИЗАЦИИ ДЕЦЕНТРАЛИЗОВАННОЙ МОДЕЛИ УПРАВЛЕНИЯ ДАННЫМИ В GRID-СИСТЕМАХ.....	65
Г.С. Краліна АНАЛІЗ МЕТОДІВ ТА МОДЕЛЕЙ ПРИЙНЯТТЯ РІШЕНЬ В СИСТЕМАХ ПЛАНУВАННЯ ТА СИТУАЦІЙНОГО УПРАВЛІННЯ....	66

К.М. Черняк	
ПОХИБКИ ВИМІРЮВАНЬ КООРДИНАТНО-ВИМІРЮВАЛЬНИХ СИСТЕМ	67
Н.А. Ковалев	
ОДИН ПОДХОД К ЛОГИЧЕСКОЙ ОБРАБОТКЕ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ОБРАЗОВ.....	68
В.Ф. Гігіняк	
ІНТЕРАКТИВНЕ РОЗВ'ЯЗАННЯ ЗАДАЧ ОПТИМАЛЬНОГО КЕРУВАННЯ НЕЛІНІЙНИМИ ОБ'ЄКТАМИ З ФАЗОВИМИ ОБМЕЖЕННЯМИ	69
Н.Ф. Халимон	
МЕТОДИКА ОЦЕНКИ КОРРЕЛЯЦИИ И РЕГРЕССИИ МЕТЕОРОЛОГИЧЕСКИХ ЭЛЕМЕНТОВ ПО РЕЗУЛЬТАТАМ МНОГОЛЕТНИХ НАБЛЮДЕНИЙ	70
О.М. Бердник	
МЕТОДОЛОГІЧНІ АСПЕКТИ КОМП'ЮТЕРНОГО МОДЕЛЮВАННЯ АЕРОГІДРОДИНАМІЧНИХ ПРОЦЕСІВ.....	71
М.В. Лупандин, И.А. Бескровная, И.М. Юрьева	
МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ПЛАНИРОВАНИЯ ИСПОЛЬЗОВАНИЯ ВОЗДУШНОГО ПРОСТРАНСТВА.....	72
А.Н. Березовский	
МОДЕЛИРОВАНИЕ ВЫЧИСЛЕНИЙ В КЛАСТЕРНЫХ СИСТЕМАХ...	73
С.М. Станко	
ВИКОРИСТАННЯ ІНТЕРВАЛЬНИХ МЕТОДІВ ДЛЯ ОПТИМІЗАЦІЇ ІНВЕСТИЦІЙНОГО ПЛАНУВАННЯ.....	74

Д.В. Ланде, д.т.н., В.Ю. Зубок
Інформаційний центр «Електронні вісти»

ЛІНГВІСТИЧНИЙ ПІДХІД ДО ВИЯВЛЕННЯ ІНФОРМАЦІЙНИХ АТАК З МЕРЕЖІ ІНТЕРНЕТ

На цей час інформація стає найважливішим стратегічним ресурсом. Щодня ми стаємо свідками процесів, коли політичні, суспільні та бізнес-конфлікти переходять в площину інформаційних протиставлень, які мають характер інформаційних операцій (ІО). Інформаційні операції визначаються як “акції, спрямовані на вплив на інформацію та інформаційні системи супротивника та захист власної інформації та інформаційних систем”.

Здійснюючи вплив на інформаційно-аналітичні системи супротивника, можна домогтися того, що особи, які приймають рішення (ОПР), зроблять неадекватні висновки, що вплине на подальший хід подій. До безпосередніх інформаційних впливів може бути віднесене розміщення в інформаційному просторі документів, що компрометують протилежну сторону, рекламу своїх переваг, перекручені дані про зовнішнє середовище, наміри тощо.

Втручання в ході ІО на інформаційно-аналітичні системи супротивника серед іншого може здійснюватися шляхами: постачання дезінформації, в тому числі підробки мережових атрибутів, електронних підписів і таке інше, що може вплинути на аналіз справжньої інформації та примусити ОПР зробити помилкові висновки; впливу на певних користувачів автоматизованих систем супротивника, що мають доступ до конфіденційної інформації; спам-атак тощо.

Для виявлення та подальшої протидії названим впливам на цей час поряд з формальними методами захисту інформаційно-комп’ютерної інфраструктури все частіше використовуються методи інтелектуального лінгвістичного аналізу, зокрема, методи глибинного аналізу текстів (Text Mining).

М.Ю. Сич

Національний авіаційний університет

СТРУКТУРА ТА МЕХАНІЗМИ РОБОТИ ІНТЕЛЕКТУАЛЬНОЇ ІНФОРМАЦІЙНО-ДОВІДКОВОЇ СИСТЕМИ «ВІКТОРІЯ»

Інтелектуальна інформаційно-довідкова система «Вікторія» представляє собою систему пошуку даних у великих сховищах інформації. Основними її завданнями є отримання та зберігання інформації на природній мові, її смислова обробка, синтез нових даних, та пошук необхідної інформації по запиту.

Вся інформація, з якою має справу система «Вікторія» надається їй на природній людській мові. Інформація може бути надана системі у простому текстовому форматі.

Система складається з окремих модулів, що використовують один одного. Це блок семантичного аналізу, блок логічних висновків, блок пошуку відповіді, блок постановки питання, блок отримання алгоритмів.

Після отримання даних система обробляє їх та зберігає у сховищі. До обробки даних відносяться операції встановлення семантичних та логіко-лінгвістичних зв'язків у них, перекладу даних з природної мови у спеціальний формат, що безпосередньо використовується системою, виділення точних фактів з інформації, синтез нових фактів з використанням методів та алгоритмів отримання логічних висновків.

Після збереження інформації у сховищі вона використовується для пошуку необхідних відомостей. Запит даних відбувається шляхом постановки системі прямого питання з питальним словом чи без нього. У разі невдалого пошуку формуються уточнюючі питання, за допомогою яких система може звернутись до користувача і отримати додаткові параметри пошуку інформації.

L.N. Badorina

National Aviation University

MODEL COUSTUCTION OF FORMAL CORRELATION

Integration processes, introduction of telecommunication means, computerizing of human activity have represented a set of new problems and tasks in scientific area which is between computer technologies and linguistics. Thanks to the current stormy development of the automated systems studies, the problem of construction of formal models describing various aspects in the subject field has become extremely important. Among them the leading position is taken by models and modes oriented on the automated evaluation of results of educational process.

It should be noted that if the construction of educational contents and integral systems in the noted area are sufficiently developed, the automation of evaluation processes is, actually, still in the initial stage. It is linked, first of all, with the circumstance, that the results of educational process appear as answers at examinations and because of that they have a naturally linguistic form. Consequently, the evaluation technology in such method gains the character of automatic (automated) comparison of naturally linguistic or fragments of texts.

Taking into account the naturally linguistic specific character of our research, the basic theoretical unit for modeling in the subject field, we have decided to assume the model of lexicographic environment (or computer-integrated lexicographic system), which was developed in a number of works.

Creating our model it is necessary to note formal correlates of linguistic constructions, which represent the essence of the subject industry, while modeling must take place both from the side of the form and from the side of the meaning. Moreover, we must take into account that a linguistic system comprises difficult hierarchy of various level complexes of units, objects and relations.

In fact, development of social knowledge persistently requires the systems of «content management». The noted problem, to our opinion, will determine the progress in communicative information technologies. This will be the key tasks for applied linguistics and linguistic technology for nearest decades.

Ю.О. Янов

Національний технічний університет України «КПІ»

МЕТОД ПОСЛІДОВНОЇ АДАПТАЦІЇ ЗАГАЛЬНИХ ПРАВИЛ ВИЯВЛЕННЯ ТА ПРЕДСТАВЛЕННЯ ІНФОРМАЦІЇ ДЛЯ ІНФОРМАЦІЙНО-АНАЛІТИЧНИХ СИСТЕМ

Кожна інформаційно-аналітична система складається із двох основних компонент: аналітичної структурованої інформації, що зберігається у базах знань, та неструктурованої, яку можна виявити у текстових документах, ділових листах, новинах. Більшість аналітичних систем сконцентрували зусилля на обробці першої компоненти, що є частковим відображенням їх інформаційного простору.

Так як якість виявлення фактів послідовно покращується з вдосконаленням алгоритмів та засобів для синтаксичного, морфологічного та семантичного аналізу, то природа виявлених знань має динамічну сутність. Якщо провести аналогію з людиною, то у певний момент людина сприймає певну подію відносно своїх поглядів та знань. Зі зміною їх, тлумачення події може змінюватись. Виходячи з цього, запропоновано метод послідовної адаптації загальних правил виявлення інформації. Даний метод відноситься до pattern-base методів [1], [2], адаптованих для української мови.

Суть даного методу полягає у можливості адаптації, тобто перебудови інформаційного простору при постійному покращенні алгоритму обробки текстової інформації. Інформація представляється у вигляді: $I=(T_s, V, L)$, де, T_s – множина текстових сегментів, V – множина правил виявлення, L – бізнес правил певної організації.

У даній роботі проводиться експеримент по виявленню знань із текстового корпусу політичних новин та використання її у інформаційно-аналітичній системі. Цей метод може бути використаний в задачах, пов'язаних з обробкою слабоструктурованих текстів.

Список літератури

1. *Dmitry Zelenko, Chinatsu Aone*, Kernel Methods for Relation Extraction, Journal of Machine Learning Research 3, 2003.
2. *Mary Califf, Raymond J. Moony*, Bottom-Up Relational Learning of Matching Rules for Information Extraction, Journal of Machine Learning Research 4, 2003.

О.В. Ругайн, к.т.н.

Національний авіаційний університет

ДИДАКТИЧНІ АСПЕКТИ КОМП'ЮТЕРНОЇ ОБРОБКИ ПРИРОДНОЇ МОВИ

Головною проблемою комп'ютерної обробки природної мови (ПМ) є різноманітні мовні неоднозначності: синтаксичні (структурні), значеннєві, відмінкові, абрєвіатурні і т.п. Тобто в разі їх наявності в мовному середовищі може бути відсутній природний зв'язок між текстом, що виражає певні знання природною мовою, і смислом, який задає семантика. Це є потенційною загрозою неадекватності тексту і семантичних даних. Тому вирішенню цієї актуальної проблеми значну увагу приділяють розробники сучасних інтелектуальних комп'ютерних систем лінгвістичного аналізу, пошукових комп'ютерних систем та мовних комп'ютерних перекладачів.

У роботі розглянуто дидактичні аспекти адаптивної комп'ютерної обробки мовних неоднозначностей ПМ у рамках існуючих технологій взаємодії ЕОМ з природною мовою за алгоритмами розробників відповідних систем.

Розробник інтелектуальної комп'ютерної системи, в якій передбачена обробка ПМ, в процесі її створення, наладки і впровадження виступає в ролі «вчителя», а відповідна система – в ролі «учня». На базі такої взаємодії можливо залучити мовно-дидактичний підхід щодо аналізу процесу навчання основним поведінковим функціям системи, а також її адаптації щодо вищезгаданих неоднозначностей ПМ.

Слід зазначити, що семантична складова процесу навчання інтелектуальної системи (тобто когнітивна) суттєво відрізняється від дидактичної. Але ж при цьому дидактична складова повинна витікати із семантичної доцільності.

Список літератури

1. *Валгина Н.С.* Теория текста: Учебное пособие. – М.: Издво МГУП «Мир книги», 1998. – 210 с.
2. *Кульчицький І. М.* Комп'ютерно-технологічні аспекти створення сучасних лексикографічних систем. – К.: НБУВ, 2002. – 57 с.

А.І. Вавіленкова

Національний авіаційний університет

СИСТЕМА АВТОМАТИЗОВАНОГО ФОРМУВАННЯ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ ТЕКСТОВОЇ ІНФОРМАЦІЇ

Внаслідок проведення аналізу різних областей штучного інтелекту, переваг та недоліків їх застосування й використання на практиці, в основу програмної реалізації методу автоматизованого формування логіко-лінгвістичних моделей текстової інформації покладено принцип роботи експертної системи.

Система автоматизованого формування логіко-лінгвістичних моделей текстової інформації (САФЛЛМ) здійснює перетворення тексту, представленого у вигляді конкретного речення, в логіко-лінгвістичну модель, тобто логічну формулу, яка являється математичним записом змісту тексту і в подальшому може використовуватися як база знань для іншої інтелектуальної системи.

Задачею САФЛЛМ є формування логіко-лінгвістичних моделей будь-якої текстової інформації, представленої у вигляді окремих речень природної мови не залежно від їх структури та тематики. Предметна область САФЛЛМ – вся природна, флективна мова. САФЛЛМ є програмною реалізацією методу автоматизованого формування логіко-лінгвістичних моделей текстової інформації. Саме ця методика робить САФЛЛМ універсальною, можливою для використання спеціалістами різних галузей. САФЛЛМ здійснює верифікацію методу, тобто є практичним підтвердженням правильності загальної ідеї методу автоматизованого формування логіко-лінгвістичних моделей текстової інформації.

Середовищем розробки інтелектуальної САФЛЛМ текстової інформації є NetBeans IDE 6.7, який надає всі засоби, необхідні для створення професійних додатків робочого середовища, корпоративних, мобільних та веб-додатків.

Мовою написання основного тексту програми обрано об'єктно-орієнтовану мову програмування Java, що обумовлено інтерактивністю, архітектурною незалежністю, можливістю роботи з базами даних, написаних в SQL та Access.

А.С. Савостеня

*Національний технічний інститут
«Харківський політехнічний інститут»*

АКТУАЛЬНІ ПИТАННЯ МОРФОЛОГІЧНОГО СИНТЕЗУ ТЕКСТІВ УКРАЇНСЬКОЇ МОВИ

Сьогодні комп'ютер навчений здійснювати чимало трудомісткої лінгвістичної роботи, у тому числі, реалізовувати етапи морфологічного синтезу і аналізу.

Методи морфологічного синтезу використовуються з метою побудови необхідної форми слова у вихідному тексті. Морфологічний синтез базується на трьох основних підходах: 1) словниковий метод синтезу; 2) алгоритмічний метод синтезу; 3) синтез на базі словника квазізакінчень.

При рішенні завдання створення системи, яка виводить відмінкові форми числівника української мови у відповідь на запит користувача, нами використовувався алгоритмічний метод морфологічного синтезу. Цей метод заснований на використанні словника основ, що представляє лексику мови, і словника закінчень, що містить граматичну інформацію.

Робота запропонованої системи ґрунтується на роботі зі словниковими статтями і їх безпосередніми компонентами (полями). Кожна словникова стаття включає: 1) заголовок словникової статті – еквівалент; 2) слова – відмінкові форми числівника.

За допомогою розробленої системи, користувач може побачити відмінкові форми числівника, який його цікавить. У разі вводу однозначного числа, якому існує еквівалент в заголовку словникової статті, вона береться без змін. Якщо вводиться двозначне число, еквівалента якому в заголовку словникової статті немає, спочатку відмінюються десятки, а потім ціле число - одиниці.

Розроблену систему можна використовувати у системах машинного перекладу для перекладу числівників. В майбутньому систему можна розвивати, наприклад розширити можливості користувача та отримувати інформацію про відмінкові форми числівників до мільйона. У такому випадку рішення задачі стає чисто технічним з використанням розробленого алгоритму.

А.С. Шевченко

Національний авіаційний університет

СТРАТЕГІЇ ІНФОРМАЦІЙНОГО ПОШУКУ

Стратегії інформаційного пошуку визначають ступінь подібності документів, що розглядаються, до пошукового запиту. Ступінь подібності визначається згідно робочої гіпотези: чим частіше пошуковий термін зустрічається в документі, тим більше «відповідним» є цей документ до пошукового запиту.

Стратегії інформаційного пошуку розробляються не тільки для визначення відповідності, але і для вирішення проблем, які пов'язані з неоднозначністю мови – один і той самий термін може позначати різні концепти (ключ в механіці означає зовсім не те, що в шифруванні), один і той же концепт може позначатись різними термінами. Стратегія інформаційного пошуку це алгоритм, який, переглядаючи набір документів (D_1, \dots, D_n), встановлює їх відповідність до пошукового запиту (ПЗ). Оскільки пошуковий термін зустрічається в документах різну кількість раз, можна говорити про різну ступінь відповідності до пошукового запиту. Алгоритм обчислює коефіцієнт відповідності (КВ) для кожного документу $KB(PZ, D_i)$, де $1 \leq i \leq n$.

Існують такі стратегії інформаційного пошуку:

- з використанням векторно-просторового представлення;
- пошук імовірності появи пошукового терміну в документі;
- з побудовою мовної моделі для кожного документу;
- з побудовою мережі припущень, яка використовується для встановлення відповідності документу до пошукового запиту;
- з булевим індексуванням, коли кожному пошуковому терміну присвоюється своя «вага», що потім враховується при побудові впорядкованих списків документів;
- з використанням не проявленого семантичного індексування;
- з побудовою нейромереж;
- з використанням продуктивних алгоритмів, коли початковий пошуковий запит «еволюційно» видозмінюється;
- з використанням нечітких множин, коли документу ставиться у відповідність нечітка множина.

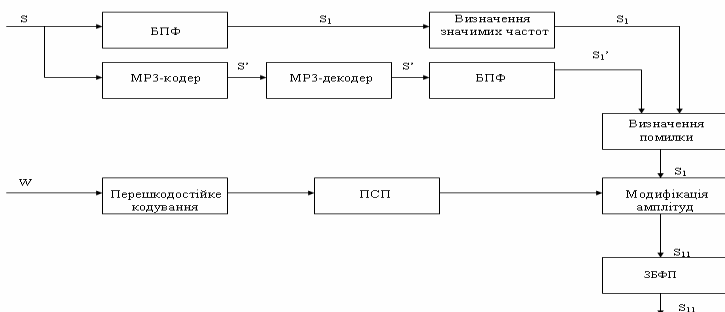
О.О. Добріна

Національний авіаційний університет

МЕТОДИ ЗАХИСТУ АВТОРСЬКОГО ПРАВА НА ЗВУКОВІ ФАЙЛИ

Сьогодні широко розповсюджене питання про виявлення плагіату. Наочно можна швидко відшукати його у тексті, але як виявити його у звуковому файлі? Щоб це реалізувати, слід просто порівняти ці файли у будь-якому звуковому редакторі – чим більше нот «співпадають», тим швидше треба робити висновки.

Одним з методів захисту від плагіату є технологія нанесення цифрових водних знаків. Вихідний звуковий сигнал представляє собою відліки імпульсно-кодової модуляції.



За допомогою швидкого перетворення Фур'є (БПФ) вихідний сигнал переводять у частотну область, де здійснюється виділення значимих частот, і вихідний сигнал поступає також на вхід Мр3-кодера, де здійснюється його стискування та обробка БПФ. Далі відбувається обчислення відносної помилки і визначення частот, в які упроваджується ЦВЗ за ще одним правилом. Сигнал ЦВЗ, представлений у вигляді 70-бітової послідовності, подається на вхід перешкодостійкого кодера, згодом генерується псевдовипадкова послідовність (ПСП) і додається в сигнал WECC. Потім виробляється модифікація амплітудних значень, вибраних за правилом частот, та перетворення сигналу шляхом зворотного БПФ у часову область. Цей метод широко використовується там, де є потрібна апаратура. Для захисту власних творів можна записати будь-який сигнал у діапазон низьких частот, що не буде сприйматися людським вухом.

К.Ю. Бабий

ООО «ГлобалЛоджик Україна»

СПОСОБ ЛОКАЛИЗАЦИИ ФУНКЦИОНАЛЬНОСТИ В ОТДЕЛЬНЫЕ МОДУЛИ

Современные технологии разработки ПО предоставляют средства для выделения логики функционирования программы в отдельные модули. Однако, не одна из этих технологий не предлагает удобного способа локализации в отдельные модули функциональности, которая должна распространяться на всю систему.

Для решения данной проблемы предлагается использовать аспектно-ориентированный подход (АОП), позволяющий выделять сквозную функциональность в отдельные модули – аспекты, и таким образом упрощать работу (отладку, модифицирование, документирование и т.д.) с компонентами программной системы и снижать сложность системы в целом.

С точки зрения АОП, в процессе разработки достаточно сложной системы программист решает две ортогональные задачи: разработка компонентов и разработка сервисов, поддерживающих взаимодействие компонентов.

Отметим, что для программирования сервисов, обеспечивающих взаимодействие объектов, нужны специальные программные средства, представляемые в рамках данного подхода. После этапа кодирования компонентов и аспектов выполняется автоматическое построение оптимизированного для выполнения (но не для просмотра и модификации) кода – слияние или интеграция (weaving).

Данная методология призвана снизить время, стоимость и сложность разработки современного ПО, в котором, как правило, можно выделить определенные части, или аспекты, отвечающие за ту или иную функциональность, реализация которой рассредоточена по коду программы, но состоит из схожих кусков кода.

Предлагаемый подход призван понизить связность внутри программной системы и, таким образом, позволяет значительно упростить сопровождение и изменение кода программы. В настоящее время АОП – единственная методология, позволяющая справиться со сложностью, присущей очень большим системам.

А.А. Остапов, А.О. Соболев

Национальный авиационный университет

СОЗДАНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ТЕСТИРОВАНИЯ КАЧЕСТВА ФУНКЦИОНИРОВАНИЯ СЕТЕЙ

Одним из ключевых понятий в IT сфере становится качество программного обеспечения. Появилась новая профессия – тестировщик и было написано множество пособий по методикам тестирования и приложений, помогающих в тестировании. И именно тут мы замечаем пробел – малое количество приложений, определяющих качество работы компьютерных сетей. Этим вопросом занимаются многие ученые, но в построении сетей нет единого и универсального правила, каждый строит сеть так, как считает нужным, и не всегда успешно. Нашей целью является создание приложения, обладающего возможностями проверить качество работы сети и указать ее сильные и слабые стороны с целью улучшения качества.

В предложенном программном комплексе предлагается реализовать удачные решения по разработке приложений: открытость исходного кода; модульность, которая позволит расширять или менять способности программы без вмешательства в основную ее часть; использование простых, удобных и современных технологий на базе Microsoft.NET; простая и понятная структура приложения, основанная на универсализации классов и четком разделении данных по назначению (рис. 1).

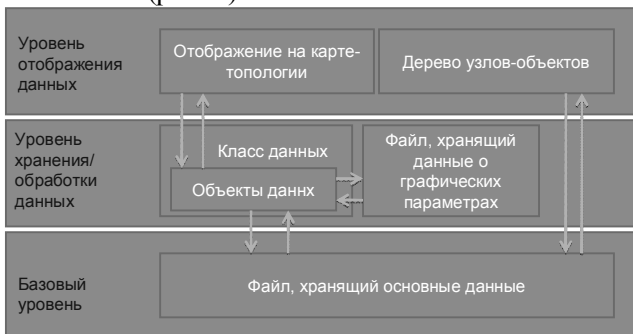


Рис. 1. Упрощенная схема работы с основными данными

М.П. Чайка

Національний авіаційний університет

ОБРАХУНОК ЕЛЕМЕНТАРНИХ ФУНКЦІЙ

В процесі вирішення багатьох практичних задач необхідно досліджувати функціональні залежності між величинами, які можна описати наближено або точно з використанням елементарних функцій (ЕФ). Необхідно відмітити, що обрахунок цих елементарних функцій займає значну частину часу всієї задачі на ЕОМ та потребує точності отриманих обрахунків.

Метою даної роботи є дослідження методів обчислень елементарних функцій з необхідною точністю.

Були виконані дослідження по методам обчислення ЕФ та конструктивним системам числення, які б забезпечували швидкість та необхідну точність обрахунку. На основі аналізу систем числення було вибрано розрядно-логарифмічне представлення даних, що забезпечує : 1) конструктивну модель виконання мультиплікативних операцій через адитивні операції, що скорочує час обчислень ; 2) точність за рахунок виключення нормалізації, округлення та використання єдиного типу даних ; 3) паралелізм обчислень.

Серед всіх методів обчислень було обрано обрахунок ЕФ з використанням ланцюгових дробів, що потребують значно менше операцій ніж, наприклад ряди Тейлора, та точність обрахунків, яку забезпечує розрядно-логарифмічне представлення даних.

На основі вибраної системи числення та методу обрахунку була розроблена модель, що підтвердила раціональність використання розрядно-логарифмічного представлення даних та ланцюгових дробів для обчислень ЕФ з необхідною точністю.

Для виконання обрахунків була використана бібліотека на мові програмування C++, що містить базовий клас спеціалізованої арифметики, основні арифметичні операції та розроблені алгоритми обрахунку основних ЕФ. Дану методику обрахунку ЕФ можна використовувати в прикладних областях інформатики та математики, які потребують точних обрахунків ЕФ. В результаті проведених досліджень були проаналізовані сучасні методи обчислень ЕФ та запропоновані нові методи вирішення даної проблеми.

І.А. Клименко, к.т.н., С.М. Біляєв
Національний авіаційний університет

РЕАЛІЗАЦІЯ СХЕМИ УПРАВЛІННЯ СТАНАМИ ТА ЗСУВАМИ НА ПЛІС

Класичним напрямом підвищення продуктивності систем управління в реальному часі є розпаралелювання обчислень [1]. У сучасних системах управління використовуються мультипроцесорні обчислювальні системи, побудовані на однотипних модулях. Збільшення числа процесорних модулів забезпечує підвищення продуктивності таких систем. Велика кількість існуючих систем, що забезпечують розв'язання задач управління різної складності, реалізовані саме за цим принципом [1]. Такі системи реалізовані на елементному базисі середньої та високої ступені інтеграції, у якості процесорних ядер звичай використовуються мікроконтролери й універсальні процесори.

Одним із сучасних напрямків розвитку мікроелектроніки є реалізація систем управління за принципом система на кристалі (*System on Chip, SoC*). Такі системи виконують функції цілого обладнання, розміщуються на одній інтегральній схемі [2] і реалізуються на програмованих логічних інтегральних схемах (ПЛІС). Мікропроцесори на одному кристалі мають ряд переваг перед звичайними мікропроцесорами середньої продуктивності: вони абсолютно синхронні з усім іншим проектом, розташованим у тій самій мікросхемі; реалізують спеціалізовані команди користувача та одержують потрібну користувачеві периферію. Реалізація мультипроцесорних систем на одному кристалі є ефективним напрямком підвищення продуктивності систем управління.

Розроблений модуль схеми управління станами та зсувами, що може бути застосований для забезпечення різних типів зсувів для арифметико-логічного пристрою RISC процесору. Пристрій розроблено на мові опису апаратних засобів *VHDL*. Розробка та моделювання пристрою виконано із застосуванням САПР *Quartus II* фірми *Altera*.

Список літератури

1. Жабин В.И. Архитектура вычислительных систем реального времени. – К.: БЕК +, 2003. – 176 с.
2. Balkan A.O., Qu G., Vishkin U. An area-efficient high-throughput hybrid interconnection network for single-chip parallel processing//Proc. IEEE /ACM Design Automation Conference (DAC). – Anaheim, CA, 2008. – P. 73 – 80/

О.С. Папук, І.О. Ніколаснко, А.М. Алексійчук
Національний авіаційний університет

ІСТОРІЯ РОЗВИТКУ ТА НАПРЯМИ КОМП'ЮТЕРНОЇ ЛІНГВІСТИКИ

Сфера комп'ютерної лінгвістики різноманітна і включає в себе такі області, як комп'ютерне моделювання спілкування, моделювання структури сюжету, гіпертекстові технології представлення тексту, машинний переклад, комп'ютерну лексографію.

Комп'ютерна лінгвістика в вузькому спектрі своєї проблематики почала розвиватися в 1960 році в рамках наукової програми «штучний інтелект». Бурхливий розвиток проблематики «обробка природної мови» приходить в 1970 –х роках і пов'язаний зі збільшенням кількості користувачів ЕВМ і появою мов програмування високого рівня. Системи обробки зв'язаних текстів досить різноманітні по структурі. Їх загальною рисою можна вважати широке використання технологій представлення знань.

Важливим напрямом комп'ютерної лінгвістики являється розробка інформаційно-пошукових систем (ІПС), що виникли в кінці 1950-х – початку 1960 років. На сучасному етапі розвитку створені ІПС безтезаурусного типу, для індексування використовують слова і словосполучення природної мови.

До області комп'ютерної лінгвістики в певній мірі можуть бути віднесені роботи в області створення гіпертекстових систем, що розглядають особливий спосіб організації тексту і навіть принципово новий вид тексту, що протилежний своїми властивостями до звичайного тексту.

В межах комп'ютерної лексикографії розробляються комп'ютерні технології складання і використання словників.

Комп'ютерне моделювання структури сюжету – ще один перспективний напрям комп'ютерної лінгвістики. Вивчення структури сюжету відносять до проблематики літературознавства.

В полі компетенції комп'ютерної лінгвістики і машинний переклад, який переживає друге народження на сучасному етапі.

Сфера комп'ютерної лінгвістики надзвичайно широка і включає різноманітні інтелектуальні технології лінгвістичного аналізу.

Ю.М. Козир

Національний авіаційний університет

ПОШУК ІНФОРМАЦІЇ ЗАДАНОЇ ТЕМАТИКИ

Задача збору інформації про Інтернет-ресурси актуальна при вирішенні таких завдань, як, наприклад, побудова тематичних каталогів типу Yahoo! або List.Ru. Близькою є задача автоматичного збору інформації про існуючі Інтернет-ресурси при створенні індексів багатоцільових пошукових систем, таких як, наприклад, Altavista, Google або Яндекс.

Для вирішення таких завдань використовуються так звані мережеві роботи – програми, які, починаючи з деякої Інтернет-сторінки, рекурсивно обходять ресурси Інтернет, витягуючи посилення на нові ресурси з отримуваних документів.

Оскільки відвідати всі Інтернет-сторінки неможливо через величезний об'єм і швидку зміну доступної в Інтернет інформації, то стратегія обходу інтелектуальних роботів визначає, які саме ресурси вдасться відвідати.

Робот, який збирає інформацію про ресурси для пошукової системи, зацікавлений у виявленні максимальної кількості різноманітних ресурсів. Подібні роботи часто використовують як оцінку «корисності» ресурсу глибину URL, тобто кількість проміжних каталогів, що згадуються в URL між ім'ям Інтернет-вузла і ім'ям самого ресурсу. Чим більша глибина, тим нижча важливість відповідного ресурсу. Такий підхід дозволяє швидко відвідати стартові і близькі до них сторінки на великому числі Інтернет-вузлів.

Сторінка, на яку посилаються багато різних сторінок в Інтернет, важливіша, ніж та, на яку мало посилань. А також, що посилення з авторитетного джерела варто оцінювати вище, ніж посилення з чієїсь персональної сторінки. Ці міркування і використовуються, наприклад, в алгоритмі мережевого робота Google.

Головною метою мережевого робота є виявлення максимальної кількості релевантних ресурсів, оцінкою яких виступає, з точки зору «корисності», його близькість до шуканої тематики. При цьому робот використовує тематичну релевантність вже виявлених сторінок.

О.О. Білик

Національний авіаційний університет

АЛГОРИТМ НЕЧІТКОГО ПОШУКУ ІНФОРМАЦІЇ

Технологія нечіткого пошуку інформації дозволяє розширювати запит близькими по написанню словами, що містяться в архіві документів, наприклад, електронної бібліотеки.

Оригінальний алгоритм здатний знайти всі лексикографічно близькі слова, що відрізняються замінами, пропусками, вставками символів та ін.

Нечіткий пошук доцільно застосовувати при пошуку слів з друкарськими помилками, а також в тих випадках, коли виникають сумніви в правильному написанні - прізвища, назви організації і тому подібне. Наприклад, запит «мультимедіа» може бути розширений словами: «сультимедіа», «іультимедіа», «мультиседіа». Якщо користувач забув, наприклад, точну назву медичного препарату «ізотрексін», то можна задати схоже слово «іподрексін»¹. Потрібні документи будуть знайдені.

Унікальні алгоритми, які реалізують нечіткий пошук, засновані на принципах асоціативного доступу до слів, що містяться в текстовому індексі повнотекстового архівного сховища документів. Як одиниці пошуку використовуються складові ланцюжки слова.

Для прискорення пошуку створюється спеціальний індекс, що містить фрагменти слів із посиланням на слова, в яких ці фрагменти зустрічаються. Алгоритм пошуку дозволяє швидко відібрати всі слова, фрагменти яких збігаються з фрагментами слова в запиті. Задаючи розмір (відсоток фрагментів, фрагментів що відрізняються, і допустимі позиційні розміщення їх в слові), можна легко регулювати точність і повноту пошуку — відбирати слова по мірі близькості до запиту.

Швидкість пошуку пропорційна логарифму від числа індексованих слів і складає менш однієї секунди при індексі в декілька мільйонів слів. Такий повнотекстовий індекс відповідає гігабайтовій розмірності повнотекстових документів).

До продуктів, що використовують технологію нечіткого пошуку інформації, відносяться: RCO for Oracle та RCO КАОТ.

И.О. Николаенко, И.В. Ефименко

Национальный авиационный университет

ИНФОРМАЦИОННЫЕ И СОДЕРЖАТЕЛЬНЫЕ ОСНОВЫ ДЕЛОВОГО ПЕРЕВОДА В ОБЛАСТИ МУЛЬТИМЕДИА

Язык – важнейшее средство человеческого общения, при помощи которого люди обмениваются мыслями и добиваются взаимного понимания. Важность изучения иностранных языков в наше время ни у кого не вызывает сомнения. Английский язык давно стал языком международного общения в сфере научных, политических, деловых контактов.

Общение людей при помощи языка осуществляется двумя путями: в устной форме и в письменной форме. Если общающиеся люди владеют разными языками, то непосредственное общение их становится невозможным. Перевод, таким образом, является важнейшим вспомогательным средством, обеспечивающим выполнение языком его коммуникативной функции в тех случаях, когда люди выражают свои мысли на разных языках.

Перевод играет большую роль в обмене информацией. В области мультимедиа, информация может представляться в самых различных формах: текст, аудио, видео, графика. Таким образом, следует сформулировать некоторые принципы, пользуясь которыми элементы мультимедийного контента должны стать понятными при их переводе на любой язык.

Первый фактор, который следует учесть – неосведомленность в области мультимедиа. При переводе с иностранного языка нужно знать сферу деятельности человека. В соответствии с этим материал подается на соответствующем уровне.

Часто идут споры о том, являются ли определенные слова непереводаемыми. Особенно остро эта проблема стоит в специфических областях перевода. Мультимедиа насчитывает ряд терминов и понятий, которые перевести нужно, точно и достоверно, чтобы даже непереводаемые темы стали понятными для аудитории.

Проведенный анализ характерных терминов в переводе на русский язык электронных материалов Internet за тематикой Multimedia, подтверждает актуальность рассматриваемой проблемы.

О.Є. Литвиненко, д.т.н.

Національний авіаційний університет

ПІДСИСТЕМА ПОРІВНЯЛЬНОГО АНАЛІЗУ ЕЛЕКТРОННИХ ТЕКСТІВ ЯК ЗАСІБ ПІДВИЩЕННЯ ЯКОСТІ ПІДГОТОВКИ ФАХІВЦІВ

В Національному авіаційному університеті створюється типова комп'ютеризована система поточного та підсумкового контролю знань студентів (ППКЗС) вищого навчального закладу.

Особливе місце у цій системі займає підсистема порівняльного аналізу електронних текстів. Її включення до складу системи пояснюється тим, що останнім часом в Україні почастишали випадки несамостійного виконання студентами письмових робіт (рефератів, курсових, дипломних проектів і робіт), передбачених навчальними планами. Все частіше студенти списують роботи з мережі Інтернет або замовляють їх у численних фірмах, які спеціалізуються у наданні послуг такого роду. Не викликає сумніву, що це явище вкрай негативно впливає на якість підготовки фахівців, оскільки, як відомо, справжні знання не передаються, а здобуваються в процесі самостійної роботи.

Для боротьби з цим явищем, яке вже набуває масового характеру, в Національному авіаційному університеті створена комп'ютеризована система порівняльного аналізу електронних текстів, призначена для виявлення текстових та змістовних збігів у студентських письмових роботах. Досвід її багаторічного використання для перевірки дипломних проектів і магістерських випускних робіт показав, що, починаючи з другого року експлуатації вказаної системи, випадки плагіату і текстових запозичень у цих проектах та роботах фактично припинилися.

Вище означене дає підставу вважати, що реалізація типової комп'ютеризованої системи ППКЗС вищого навчального закладу з підсистемою порівняльного аналізу електронних текстів сприятиме підвищенню якості підготовки фахівців.

УДК 004.912

Т.В. Шепелева, к.фіз.-мат.н., Г.С. Краліна
Національний авіаційний університет

ДОСЛІДЖЕННЯ І РОЗРОБКА ЕФЕКТИВНИХ АЛГОРИТМІВ ТОКЕНІЗАЦІЇ

Сучасні технології попередньої обробки електронних текстів базуються на багаторівневому представленні природної мови. Компоненти, що складають мовну модель, – лінгвістичні процесори, які один за одним обробляють вхідний текст, – вхід одного процесора є виходом іншого. Як правило, виділяються наступні компоненти процесу попередньої обробки електронних текстів:

– токенізація – графематичний аналіз – виділення слів, цифрових комплексів, формул тощо;

– морфологічний аналіз – побудова морфологічної інтерпретації слів вхідного тексту;

– синтаксичний аналіз – побудова дерева залежностей речення в цілому;

– семантичний аналіз – побудова семантичного графа тексту.

Графематичний аналіз є початковим етапом обробки тексту, в ході якого визначаються елементи граматичної структури (слова, знаки пунктуації, числа, скорочення тощо). Результатом такого аналізу має бути інформація, що є необхідною для подальшої обробки морфологічним та синтаксичним процесорами.

В роботі виконано огляд сучасних алгоритмів токенізації з точки зору їх ефективності, розглянута постановка задачі графематичного аналізу в термінах теорії формальних мов, обрана графематична модель російської та української мови, на підставі якої виконана алгоритмічна і програмна реалізація модуля для графематичного аналізу.

Запропоновано тести для аналізу ефективності алгоритмів графематичного аналізу електронних текстів.

М.В. Лупандін, к.т.н.,
І.А. Безкровна, І.М. Юр'єва
Національний авіаційний університет

РОЗРОБКА АЛГОРИТМІВ СЕМАНТИЧНОЇ СТРУКТУРИЗАЦІЇ ДАНИХ

Проблеми попередньої обробки електронних текстів в програмних системах відповідного призначення, в тому числі у пошукових системах, не втрачають актуальності протягом останніх років. Аналіз технологій попередньої обробки електронних текстів показав, що сучасні механізми структуризації даних мають передбачати можливість аналізу даних, що представлені в будь-якому форматі – у вигляді Web-сторінок, документів Microsoft Word, файлів Adobe Acrobat тощо. Розв'язання цієї задачі суттєво полегшує пошук та доступ до інформаційного змісту даних.

Задача пошуку текстової інформації полягає в знаходженні мінімальних смислових одиниць тексту, які є релевантними запиту. Знайдені одиниці мають відповідати вимогам повноти і точності. Під релевантністю розуміється деяка бінарна функція, вхідними параметрами якої є запит користувача і чергова одиниця тексту, що підлягає аналізу. За інформаційну одиницю тексту може братися документ, абзац, речення або інші фрагменти тексту. Проблема обчислення релевантності є центральною в задачі пошуку текстової інформації.

Виконано порівняльний аналіз лінгвістичного і статистичного підходів до попередньої обробки електронних текстів. З порівняння сучасних статистичних та нестатистичних підходів до аналізу текстів на природній мові можна зробити висновки, що найбільш ефективними будуть алгоритми, які поєднують в собі швидкість і незалежність від мови статистичних алгоритмів, з високою якістю обробки семантичного підходу.

В роботі розроблено ефективні алгоритми семантичної структуризації даних. Запропоновано тести для аналізу ефективності алгоритмів семантичної структуризації електронних текстів.

О.М. Дишлок, О.П. Нечипорук
Національний авіаційний університет

РОЗРОБКА МЕТОДІВ ТА АЛГОРИТМІВ ІНДЕКСАЦІЇ ЕЛЕКТРОНИХ ТЕКСТІВ

Розглянуто загальну постановку задачі класифікації і базовий підхід до індексації електронних текстів.

Виконано аналітичний огляд сучасних алгоритмів індексації, розглянуті оцінки їх ефективності. Окремо проаналізовано алгоритми індексації для документів, що зберігаються в графічних форматах TIFF, JPEG и PDF.

Особлива увага була приділена методу кореляційної індексації текстової і графічної інформації, який дозволяє забезпечити високу релевантність пошуку та істотно скоротити його часові витрати. Метод кореляційної індексації при створенні індексу архіву електронних документів враховує кореляції між словами або ASCII символами. Облік кореляцій дозволяє при індексації автоматично виявляти термінологію (як слова, так і послідовності слів), характерну для документів, записаних в архів. Такий підхід дозволив розробити алгоритм пошуку схожих документів, при якому як пошуковий запит може використовуватися повна сторінка тексту. Пошук схожих документів дає можливість не лише шукати необхідну інформацію, а й аналізувати та класифікувати цю інформацію. В разі індексації текстової інформації з врахуванням кореляцій між ASCII символами пошук схожих документів є мовно-незалежним.

Створено декілька алгоритмів, на підставі яких буде виконана програмна реалізація індексації електронних текстів.

Спроектовано структуру програмного модуля індексації електронних текстів.

Запропоновано тести для аналізу ефективності розроблених алгоритмів.

В.И. Надточий, к.т.н.,

Е.Д. Надточий, к.фил.н.

Национальный авиационный университет

НПУ им .М.П. Драгоманова

АЛГОРИТМИЗАЦИЯ ПРОЦЕССА АНАЛИЗА ЯЗЫКОВЫХ ЯВЛЕНИЙ

Исследование приемов формализации правил построения языковых структур базируется на основе методов структурной лингвистики, которые позволяют анализировать различные явления как естественного, так и искусственного языков. Для построения современных интеллектуальных информационных систем и систем поиска информации широко используются языковые модели орфографии. Теоретические и практические разработки в направлении создания и применения эффективных методик алгоритмизации орфографии, к сожалению, остаются на периферии интересов ученых-языковедов.

В докладе рассматриваются проблемы лингвистического анализа и синтеза, орфографических правил в виде алгоритмов. Процесс создания алгоритмов базируется на основе модельного подхода к способу описания сложных явлений как совокупности простых, единичных элементов. Алгоритмическая модель включает небольшую по объему базу данных и правил их кантоминации для описания максимального количества языковых объектов.

Важным условием построения алгоритмической модели является равновесие между степенью элементарности исходных утверждений и их количеством, с одной стороны, и сложностью порождаемых единиц высших уровней, с другой стороны. Простота модели определяется равнодействующей этих двух факторов.

Комплекс языковых явлений исследован в части методики алгоритмизации построения моделей орфографических явлений на материале русского и украинского языков. Дальнейшие направления исследования алгоритмических моделей связаны с использованием их для совершенствования алгоритмов анализа запросов в системах автоматизированного поиска данных.

Д.Ю. Катюха

Національний авіаційний університет

ОБ'ЄКТНИЙ ПІДХІД ДО КОМП'ЮТЕРНОГО АНАЛІЗУ ТЕКСТІВ ЗА ЗМІСТОМ

Порівняльний аналіз текстів машиною (комп'ютером) в наш час є досить складним завданням, адже будь-яка природна мова є надзвичайно складною і містить безліч тонких, не завжди помітних зв'язків, що надає складності її аналізу для машини. Також одним з факторів, що ускладнює розвиток комп'ютерного аналізу текстів, є можливість одного і того ж слова/фрази приймати різні (а іноді і зовсім протилежні) значення; можливість не однозначного трактувати те чи інше твердження; відсутність чіткої структури речення (особливо в розмовній мові). Ще одним важливим фактором, що підвищує складність аналізу є зміна властивостей об'єктів та середовища в часі.

Ідея даного методу полягає в об'єктному підході до аналізу тексту комп'ютером. Кожне речення описує якийсь об'єкт (дію, що теж є об'єктом) і набір властивостей даного об'єкта (і/або дію, що виконується об'єктом, або над об'єктом) та/або зв'язки даного об'єкта з іншими (наприклад, порівняння: "об'єкт1 світліший за об'єкт2"). Таким чином утвориться певна структура об'єктів, пов'язаних між собою, які матимуть набір властивостей, що описує реальний об'єкт. Порівняння і визначення змісту здійснюватиметься за базою синонімів, де слова мають бути представлені у вигляді синонімічних рядів. Причому різні синонімічні ряди можуть бути пов'язані між собою ступенем близькості значення.

Серед недоліків об'єктного підходу до комп'ютерного аналізу тексту за змістом є перш за все складність його реалізації, оскільки доведеться опрацювати великі масиви даних, що пов'язані між собою.

ИСПОЛЬЗОВАНИЕ КЛЮЧЕВЫХ СЛОВ И ВНЕШНИХ ТЕМАТИЧЕСКИХ ССЫЛОК ДЛЯ ПОИСКОВОЙ ОПТИМИЗАЦИИ САЙТОВ

При индексации web-страниц, поисковые роботы придают большое значение ключевым словам и ссылкам на сайты смежной тематики. Чтобы сайт выдавался в числе первых при поисковом запросе, надо правильно продумать содержание web-страниц.

Нужно упорядочить содержание каждой страницы сайта и разбить их на разделы. Разделы необходимо озаглавить ключевыми словами, которые относятся к теме сайта. При индексации поисковый робот учитывает выделенные слова.

Необходимо внедрить ключевые слова в тело самого текста. При этом не стоит ими злоупотреблять. Во-первых, на определенном этапе робот перестает учитывать повторяющиеся слова, и их наличие не принесет никакого положительного эффекта. Во-вторых, если текст будет перенасыщен одинаковыми словами, пользователи не захотят возвращаться на такой сайт. Для оптимальной индексации в самых популярных поисковых машинах процент содержания ключевых слов в тексте составляет около 10-12 % от общего количества слов. Также нужно заполнить атрибуты description и keywords meta-тэга. Хотя сейчас из-за злоупотребления данными тэгами, их вес для поисковых машин незначителен, но все же они влияют на место в рейтинге поисковых машин. Также атрибут description имеет очень большую значимость при выдаче результата поиска: его содержание выводится как комментарий к ссылке на страницу и по нему пользователь судит о соответствии ссылки его запросу.

На рейтинг страницы влияет наличие внешних ссылок на популярные сайты смежной тематики. Еще большее влияние на рейтинг оказывают ссылки таких сайтов на ваш сайт. Рекомендованное количество внешних ссылок – не меньше 50. При этом стоит избегать ссылки на сайты, которые не относятся к теме и имеют низкую популярность – они понизят ваш собственный рейтинг.

УДК 681.518

Є.Б. Артамонов

Національний авіаційний університет

КОНЦЕПЦІЇ СТВОРЕННЯ ЕЛЕКТРОННИХ ПІДРУЧНИКІВ З МОЖЛИВІСТЮ АВТОМАТИЧНОГО НАЛАШТУВАННЯ ГЛИБИНИ ПРЕДСТАВЛЕННЯ ІНФОРМАЦІЇ

Розробка електронних підручників зі змістом, який автоматично змінюється за рівнем глибини представлення інформації, можлива лише у спеціальному середовищі, яке повинне забезпечувати виконання наступних задач:

- 1) створення самої книги;
- 2) визначення зв'язків між текстом книги та його змістом;
- 3) фільтрування поданої інформації на основі зацікавленості користувача питаннями, які висвітлені в окремих підрозділах підручника.

Об'єктами досліджень у даному напрямку є механізми автоматичного і автоматизованого аналізу змісту тексту та алгоритми для створення, редагування й аналізу внутрішніх зв'язків між модулями тексту.

Для реалізації подібної системи необхідно послідовно вирішити наступні задачі:

– розробити загальну концепцію програмного середовища для реалізації подібних електронних підручників;

– вивчити залежності прямих і вторинних ознак зацікавленості текстом;

– розробити алгоритми та програмні засоби оцінювання зацікавленості змістом тексту;

– налагодити гнучкий механізм для відстеження зацікавленості читача, що дозволить своєчасно надавати доступ до потрібної інформації;

– розробити механізми коректного зв'язування елементів тексту, яке дозволить подавати інформацію як у повній формі, так і у реферативному вигляді, при цьому реферативна форма повинна максимально передавати загальний зміст і, за необхідності, розкривати окремі ділянки тексту, які можуть зацікавити читача.

Т.Г. Куцар

Національний авіаційний університет

ВИЛУЧЕННЯ ФАКТОГРАФІЧНОЇ ІНФОРМАЦІЇ В ТЕКСТОВИХ ДОКУМЕНТАХ

Тема дослідження наукової роботи – вилучення фактографічної інформації з текстових документів особливого стилю, до яких можна віднести біографії, протоколи та інші документи, призначення яких полягає в лаконічній передачі сукупності фактів про деякі об'єкти. Ключовою особливістю таких текстів є висока щільність зв'язків між словами, що не виражаються граматичними засобами, а поєднуються анафоричними зв'язками.

Метою роботи є аналіз встановлення анафоричних зв'язків в текстових документах.

Постановка задачі комп'ютерного аналізу текстів вимагає розпізнавання і класифікації описаних у них фактів з подальшим перетворенням інформації в записи БД. Результати аналізу тексту описуються фреймовою моделлю: кожен факт формує запис у відповідній таблиці, ім'я якої визначає тип факту, імена стовпців – ролі фігурантів факту, а значення в полях таблиці – імена учасників-фігурантів факту у відповідних ролях. Для вирішення поставленої задачі були використані наступні засоби:

- модуль розпізнавання особливих текстових конструкцій (реєстраційних даних) на підставі шаблонів, написаних на спеціальній формальній мові;

- синтаксичний аналізатор, що визначає лексико-граматичні характеристики елементів тексту і перетворює текст кожного речення в семантичну мережу;

- правила розв'язання кореферентності власних назв, у тому числі й анафоричних зв'язків, які дозволяють ототожнити займенники, а також повні і короткі позначення персон і організацій;

- правила встановлення анафоричного зв'язку між вільною синтаксичною або семантичною валентністю, яка відповідає опущеному в реченні фігуранта факту, і згадкою одного з можливих референтів.

Схема виділення фактографічної інформації практично може бути реалізована в програмному комплексі.

С.М. Денисенко

Національний авіаційний університет

АНАЛІЗ ПАРАМЕТРІВ МОВИ У КОМП'ЮТЕРНИХ СИСТЕМАХ ОБРОБКИ СИГНАЛІВ

При створенні автоматизованих навчальних систем поряд із формуванням та представленням аудіо контенту необхідно значну увагу приділити проектуванню та розробці мовного діалогу між особою, що навчається, та комп'ютером.

Мовне введення та виведення інформації передбачає якісно нове спілкування між людиною та машиною і покликане покращити взаємодію та підвищити процес оволодіння новими знаннями. Розробки щодо згаданого аспекту велися ще з 70-х рр. спеціалістами галузі кібернетики, розпізнання образів та мови. Сьогодні існують експертні системи навчального призначення, здатні як «отримати і зрозуміти» мовне звернення людини, так і відповісти на нього. Автоматизоване розпізнання мови – це віднесення мовного сигналу до вибраного людиною класу. Існує по фонемне розпізнання мови, розпізнання окремо вимовлених слів та розпізнання злитної мови.

Особливо важливим є розпізнання аудіосигналу, що здійснюється на основі параметрів мови, які відображають індивідуальні особливості голосу, а саме: спектральні, формантні, часові та амплітудні характеристики мовних сигналів. Даним параметрам властиві такі ознаки: легкість у вимірюванні, мала залежність від завад, стабільність у часі. Слід зважати і на те, що сприйняття мовних сигналів органом слуху є дискретним. Це необхідно враховувати при виборі формату обробки (стиснення) оцифрованої інформації.

Не менш важливим є зв'язок фізичних параметрів мови та емоційного супроводження. Тобто, зважаючи на те, що потік мовної інформації включає такі компоненти, як змістовність, мелодика, індивідуальне забарвлення та динаміка, необхідно враховувати інтегральні властивості мови при комп'ютерній обробці.

Розробки даної проблеми матимуть широке практичне застосування в адаптивних навчальних системах, зокрема, у дистанційній освіті, коли особі, що навчається, необхідна швидка експертна оцінка її усної відповіді при виконанні відповідного навчального завдання.

УДК 004.01(082)

Д.В. Ланде, д.т.н., В.В.Жигало
Інформаційний центр «Електронні вісті»
НТУУ «КПІ»

МЕТАПОИСКОВАЯ СИСТЕМА PDF SCIENCE SEARCH

В настоящее время в веб-пространстве содержится большое количество документальных ресурсов, представленных в формате PDF. Вместе с тем, при поиске необходимой документации с помощью традиционных сетевых информационно-поисковых систем (ИПС) пользователь постоянно сталкивается с проблемами, связанными с плохой доступностью целевой информации.

Система PDF Science Search (PDFSS) основана на использовании существующих сетевых ИПС, таких как scholar.google.com, yandex.ru, rambler.ru. Основная идея данной метапоисковой системы состоит в том, чтобы находить в Сети pdf-файлы без сопровождающего их информационного шума или рекламы (до настоящего времени такой системы не существовало). Особенностью PDFSS является то, что она полностью направлена на поиск доступных пользователю pdf-файлов, с возможностью фильтрации платных ресурсов, текстовых описаний, всего, кроме самих файлов.

PDFSS использует другие ИПС для наполнения собственной базы данных (БД, кеша системы). После того, как пользователь вводит запрос система отправляет его каждой из выбранных ИПС, обрабатывает полученные данные, дополняет информацией, такой как, наличие файла в БД, существование файла по данному адресу в Интернет и выявляет подобные файлы из других источников. Система собирает новые адреса файлов, после чего в фоновом режиме производит их скачивание и занесение метаинформации в БД (наличие файла по адресу, доступность сервера). Кроме того, происходит индексация и сохранение текстовых образов pdf-файлов в кеше системы.

Система PDFSS содержит встроенную ИПС, обеспечивающую поиск в кеше системы как по названиям файлов, так и по контексту. Данная ИПС обеспечивает ранжирование результатов поиска, а также просмотр текстовых образов найденных документов. Система PDFSS доступна по адресу <http://chaos.in.ua/>.

УДК 681.518

В.В. Ковальчук

*Національний технічний інститут
«Харківський політехнічний інститут»*

ЗАСОБИ ПІДВИЩЕННЯ РЕЛЕВАНТНОСТІ ПОШУКУ В ІНФОРМАЦІЙНИХ СИСТЕМАХ НАУКОВИХ БІБЛІОТЕК

Традиційний довідково-бібліографічний апарат, який включає системи каталогів і картотек та довідково-бібліографічні фонди, незважаючи на недоліки, є більш інтелектуальною інформаційно-пошуковою системою, ніж пошукові системи існуючих електронних каталогів.

Багатоаспектне відображення бібліотечного фонду в електронному каталозі дає можливість проводити оперативний пошук інформації за будь-яким елементом бази даних, у тому числі і за тими, що не можуть бути використані при пошуці в традиційних каталогах, наприклад, за індексом ISBN, ISSN, роком видання і будь-якою комбінацією ознак з урахуванням словоформ.

Поряд з традиційним способом тематичного подання знань для формування повноцінного лінгвістичного забезпечення сучасної інформаційної служби бажано застосовувати комплекс додаткових лінгвістичних засобів більш об'єктивного та розширеного подання змісту електронних документів на основі досягнень комп'ютерної лінгвістики та методів адаптації пошуку.

Для досягнення якісно нового результату функціонування інформаційно-пошукових систем наукових бібліотек необхідним є розвиток та впровадження нової парадигми інформаційного пошуку, які передбачають ранжирування результатів пошуку на основі використання різнопланових статистичних даних для оцінки релевантності документів і розробку адаптивного інтерфейсу користувача.

З розвитком обсягів впровадження комп'ютерних технологій в бібліотечну діяльність принципово важливою стає організація інформаційних зв'язків між усіма автоматизованими бібліотечними процесами не лише в межах однієї бібліотеки, а й у рамках бібліотечних об'єднань.

УДК 004.01(082)

Е.Ю. Романец, А.В. Лищинский
Национальный авиационный университет

ТЕОРИЯ И ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО ПОИСКА

Информационно-поисковая система (ИПС) – система, выполняющая функции: хранения больших объемов информации; быстрого поиска требуемой информации; добавления, удаления и изменения хранимой информации; вывода информации в удобном для человека виде.

Технологические характеристики: скорость обработки запросов; полнота охвата ресурсов; вероятность получения ответа от системы; нахождение документов, подобных найденным; возможность уточнения запросов; возможность применения разных критериев ранжирования выдачи; подключение переводчиков.

Недостатки традиционных ИПС:

- недостаточная оперативность;
- зависимость от выбора источников;
- отсутствие средств уведомления о нахождении новой информации;
- невозможность работы с мультимедиа;
- недостаточная защита данных;
- слабо развитые средства обобщения данных.

Булева модель является классической и широко используемой моделью представления информации, базирующейся на теории множеств, и, следовательно, и моделью информационного поиска, базирующейся на математической логике. Векторно-пространственная (Vector Space Model) модель является классической алгебраической моделью. В рамках этой модели документ описывается вектором в некотором евклидовом пространстве. Вероятностная модель поиска базируется на теоретических подходах байесовских условных вероятностей. Основным подходом вероятностной модели является вероятностная оценка веса термов в документе. С другой стороны, в качестве оценки соответствия документа запросу используется вероятность того, что пользователь признает документ релевантным.

УДК 004.01(082)

Т.О. Слухай

Національний авіаційний університет

МЕТА І ЗАВДАННЯ ДОСЛІДЖЕННЯ ТЕХНОЛОГІЙ ІНФОРМАЦІЙНОГО ПОШУКУ

Метою роботи є підвищення ефективності інформаційного пошуку в Інтернеті шляхом створення методів і засобів його адаптації до інформаційних потреб користувачів за допомогою моделі користувача.

Реалізація мети передбачає вирішення таких завдань:

– аналіз існуючих методів і засобів моделювання користувачів в інформаційних системах;

– розробка моделі адаптивного інформаційного пошуку на основі моделі користувача;

– розробка моделі користувача інформаційно-пошукової системи, що враховує контекст інформаційного пошуку;

– розробка методів автоматичного формування моделі користувача і модифікації запитів на її основі з чіткою репрезентацією контексту пошуку;

– розробка методів автоматичного формування стереотипів користувачів шляхом динамічного об'єднання інформації про сеанси пошуку в категорії.

Об'єкт дослідження становить процес інформаційного пошуку в Інтернеті.

Предметом дослідження є методи та засоби адаптивного інформаційного пошуку на основі моделі користувача.

Методи дослідження ґрунтуються на використанні теорії інформаційного пошуку, теорії моделювання і методів системного аналізу – для побудови формальної контекстної моделі користувача; теорії машинного навчання – для розробки методів автоматичного формування контекстної моделі користувача і стереотипів користувачів; методів дедуктивного виведення – для автоматичного розширення запиту користувача; теорії агентних систем – для розробки архітектури і моделі роботи інформаційно-пошукової мульти-агентної системи з контекстною моделлю користувача.

Д.А. Бурко

Національний авіаційний університет

ОПТИМІЗАЦІЯ МОДУЛЯ ПОШУКУ ДАНИХ В ІНФОРМАЦІЙНИХ СИСТЕМАХ

В основі сучасних інформаційних систем лежать програмні модулі, побудовані на основі реалізації пошуку інформації у базах даних. Головне завдання таких систем полягає у створенні запиту за допомогою мови SQL (Structured Query Language) з подальшою візуалізацією результату.

Під час розробки програмних модулів варто застосувати метод централізації функціональності за принципом «чорної скриньки» або системи «вхід – вихід».

На вхід будь-якої пошукової системи подається низка параметрів, які дозволяють виділити тип інформації, необхідний для підготовки оптимального SQL-запиту. Цікаво відзначити, що модуль являє собою складну ієрархічну структуру. Преференційним можна вважати застосування об'єктно-орієнтовного підходу, що відрізняється високим рівнем чутливості до модифікації інформаційної системи. Сам програмний модуль пропонуємо представити у вигляді кількох атомарних пошуків (пошук даних над певним аутентичним типом інформації), семантичного аналізатора, внутрішньої статистики модуля та системи прийняття рішень по вибору доцільної пошукової стратегії.

На рівні атомарного пошуку запиту обробляються швидко, але об'єднання їх результатів не завжди видасть найкращу стратегію пошуку. Тому її вибір має бути реалізований на рівні внутрішнього програмного коду. Різниця між стратегіями базується на статистиці пошукового модуля та вбудованих описаних правилах.

Перевагою централізації логіки у пошукових системах є її компактність. Без сумніву, створення подібного модуля потребуватиме деяких затрат часу. Проте результат представлення даних та спрощення подальшої розробки і підтримки системи в даному випадку є серйозною перевагою. На такому підході позначається тенденція до постійного вдосконалення інформаційних систем.

УДК 004.01(082)

К.А. Мацуева

Национальный авиационный университет

ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО ПОИСКА (ТЕМАТИЧЕСКИЙ ПОИСК)

Развитие индустрии систем электронного документооборота, сопровождающееся ростом массивов обрабатываемых полнотекстовых документов, требует новых средств организации доступа к информации, многие из которых следует отнести к разряду систем искусственного интеллекта.

Основной задачей, возникающей при работе с полнотекстовыми базами данных, является задача поиска документов по их содержанию. Однако, ставшие традиционными средства контекстного поиска по вхождению слов в документ зачастую не обеспечивают адекватного выбора информации по запросу пользователя.

Основная проблема заключается в сложности точной формулировки запроса - подбора ключевых слов, которые предстоит искать в телах документов. Это может быть связано с рядом причин, как то: недостаточным знанием пользователем терминологии предметной области, трудностью определения границ своих интересов, наличием в языке многозначных и синонимичных слов, и даже орфографическими ошибками в написании искомых слов, которые могут встречаться как в текстах, так и в самом запросе.

Отдельную сферу приложения технологий обработки текста занимает широкий класс задач информационного поиска, относимых к разряду аналитических. Это задачи, требующие "раскопки" скрытых взаимосвязей между целевыми объектами, выявления ключевых факторов, влияющих на развитие ситуации, прогнозирование тенденций ее развития и т.п.

Над решением указанных проблем работает наше подразделение RCO Research Group. Результат многолетней работы воплотился в ряду технологий, основанных на автоматическом анализе содержания текстовых документов, применение которых призвано облегчить процессы поиска и эксплицирования скрытой информации. Часть технологий реализована в наших "коробочных" продуктах, другая часть находит внедрение в заказных решениях.

УДК 004.416.2(043.2)

Н.П. Кадет, М.В. Сорока

Національний авіаційний університет

ВИКОРИСТАННЯ СЕМАНТИЧНОГО ПОШУКУ В ІНФОРМАЦІЙНИХ СИСТЕМАХ

Семантичний пошук визначається як вид автоматизованого повнотекстового інформаційного пошуку з урахуванням смислового змісту слів і словосполучень запиту користувача і пропозицій текстів проіндексованих інформаційних ресурсів [1].

Виникають потреби при розробці системи переведення російського (українського) тексту в кодовану мову. У разі перекладу із словесного на кодовану мову проблеми зняття омонімії відрізняються від подібних завдань при перекладі з однієї словесної мови на іншу. Методики, що базуються на використанні тезаурусів застосовні лише частково. У зв'язку з тим, що коди передають не значення слів, а концепти, що часто описуються групою слів, запропонований метод, що розширює положення.

Однією з наймогутніших сучасних систем, що використовує технології семантичного пошуку, є *RetrievalWare* компанії *Convera* (раніше *Excalibur*), що позиціонує свою систему як «першу платформу по витяганню знань» [2]. Основою семантичного пошуку в *RetrievalWare* є використання семантичних мереж, що описують сенс слів мови і зв'язку між поняттями, що позначаються ними.

Запропонована методика семантичного пошуку дозволяє відбирати близькі за контекстом документи, навіть якщо вони належать до різних предметних областей. Пропонується, що запропонована методика може бути більш корисною при перетворенні *Internet*-мережі в Семантичну мережу. В результаті проведених експериментів отримана математична модель запиту [2]:

$$Y_{2i} = 0.517 + 0.157A + 0.207AB - 0.264AC + 0.122BC - 0.246ABC$$

Для створення програми переведення тексту в символи, необхідно використовувати засоби семантичної обробки тексту, враховуючи весь попередній та локальний контекст.

Список літератури

1. http://www.stocona.ru/glossary/semantic_search.aspx.
2. Т. Гаврилова Онтологія знань. - Труды МНПК, 2001.

УДК 004.9

**А.Ю. Михайлюк, О.В. Пилипчук,
М.В. Сніжко, В.П.Тарасенко**
НТУУ «КПІ», КМПУ імені Б.Д. Грінченка

СПОСІБ СТРУКТУРНО-АЛГОРИТМІЧНОЇ ОРГАНІЗАЦІЇ ІНТЕЛЕКТУАЛЬНОГО РЕДАКТОРА ЗАПИТУ НА КВАЗІСЕМАНТИЧНИЙ ПОШУК

Важливим чинником успішної професійної діяльності в умовах суспільства знань є наявність засобів оперативного отримання вичерпної довідкової інформації з актуального питання. У глобальний інформаційний простір сьогодні інтегровано велику кількість інформаційно-пошукових систем. Однак, функціональність сучасних засобів пошуку інформації далеко не повністю задовольняє все більш вибагливі потреби користувачів. Однією з найважливіших вад сучасних «пошуковиків» є відсутність можливості ефективного семантичного пошуку інформації, передусім текстової. Справа в тому, що переважна більшість розроблених на сьогоднішній день методів семантичного пошуку, по-перше, є обчислювально складними, вимагають наявності адекватної моделі природної мови (тобто надзвичайно об'ємного і складного у плані реалізації службового лінгвістичного інформаційного ресурсу), принципово не дозволяють ідентифікувати всі відтінки змісту з огляду на контекст.

У доповіді пропонується оригінальний інструментарій для реалізації так званого квазісемантичного пошуку. Побудований на кшталт експертної системи інтелектуальний редактор запиту дозволяє у інтерактивному режимі сформувати оптимізований щодо семантики пошуковий критерій, власне пошук за яким буде зводитись практично до класичного повнотекстового. Таким чином, розв'язання неформалізованої на сьогоднішній день до кінця задачі виявлення семантичної відповідності бере на себе користувач при інформаційній підтримці інтелектуального програмного інструменту, який може бути виконано у формі програмної надбудови над стандартною інформаційно-пошуковою системою.

Доповідь містить детальний опис структурно-алгоритмічної організації інтелектуального редактора запиту, зокрема спеціальної бізнес-логіки, оригінального підходу до реалізації онтологічної бази знань та способу ітераційної інтерпретації її вмісту тощо.

УДК 004.45(043.2)

**М.А. Мелешко, к.т.н., С.М.Денисенко,
Ю.В. Помазков, к.т.н.**

*Національний авіаційний університет
ТОВ «Інформаційні бізнесові системи і телекомунікації»*

МОДЕЛЬ СУБ'ЄКТНОГО ОТОЧЕННЯ ВІДКРИТОЇ АРХІВНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ

Розвиток освітніх послуг в умовах інтеграції України в європейську економічну та освітню системи потребує створення інформаційних ресурсів в електронному форматі, зберігання яких може бути у відкритих архівних інформаційних системах (ВАІС) електронних бібліотек (ЕБ). Пропонується модель ВАІС, яка визначає правила довгострокового зберігання (консервації) та доступу до інформації. Спрощена структура ЕБ наступна:

- виробники, люди або клієнтські системи, які постачають;
- інформацію, що зберігається;
- цифрове сховище ВАІС, яке складається із систем та персоналу, що забезпечують збереження інформації та надання професійних послуг користувачам;
- адміністратор – функція, яка встановлює правила роботи цифрового архіву;
- користувачі – це люди, або клієнтські системи, що ведуть пошук та придбання інформації.

Розглядається модель функціонально-компонентного складу процесів створення та функціонування ЕБ, основним компонентом якого являється цифрове захоплення (ЦЗ). ЦЗ забезпечує послуги та функції прийому інформаційних вхідних пакетів (ІВП) від виробника та готує контент для завантаження та обробки всередині ВАІС. Режим захоплення забезпечує прийом ІВП, контроль якості та генерацію інформаційних архівних пакетів (ІАП), які відповідають заданим форматам і стандартам.

В роботі подається описова характеристика функцій, які виконує ЦБ: архівне збереження, керування даними, адміністрування, планування цифрової консервації, доступ до ВАІС. Розглядаються функціональні схеми процесу вводу даних в архів, управління даними та метаданими, цифрового архівного збереження та доступу до цифрового архіву.

УДК 004.424.4:025.135(043.2)

О.В. Козлова

Національний авіаційний університет

МЕТОДИ ІНФОРМАЦІЙНОГО ПОШУКУ В ЕЛЕКТРОННИХ БІБЛІОТЕКАХ

Електронні ресурси бібліотеки обов'язково включають, як системну компоненту, пошуковий апарат (системи метаданих, в тому числі бібліографічні, реферативні, інші бази даних та лінгвістичне забезпечення) [1].

Пошукова система (ПС) має бути уніфікованою для усіх різновидів інформації: текстової, аудіо та відео, графічної, мультимедійної в широкому розумінні. Бажано, щоб ПС мала:

- інтуїтивно сприятливий інтерфейс пошуку;
- можливості пошуку з використанням лише стандартного програмного забезпечення;
- пошук за метаданими цифрового об'єкту (параметричний пошук), наприклад, за інвентарним номером, назвою видавництва;
- пошук за ключовим словом (повнотекстовий пошук);
- інтегрований пошук, як комбінація відомих методів.

В рамках роботи студентського навчально-науково-дослідного центру "MULTIMEDIA" кафедри комп'ютерних мультимедійних технологій створюється кафедральна електронна бібліотека «Мультимедійні навчальні ресурси». Розробка передбачає чотири типи пошуку: за темою, назвою, розширений та останні завантаження. Користувач має можливість самостійно обирати той чи інший пошук. Крім того, в цій системі передбачений пункт під назвою «Бажане». Він необхідний для того, щоб користувач зміг повідомити про інформаційний ресурс, якого в даний момент немає в базі даних бібліотеки, але який користувач хотів би бачити. Це дозволить розширити можливості самостійної роботи студентів.

Список літератури

1. *Баркова О.В.* Модельний підхід у дослідженні технологічного забезпечення електронної бібліотеки // Реєстрація, зберігання і обробка даних. – 2003. – Т. 5, № 3. – С. 81-90.

УДК 004.4'414(043.2)

Н. Л. Соколова

Національний авіаційний університет

СЕМАНТИЧНИЙ МЕХАНІЗМ В ТЕХНОЛОГІЯХ ІНФОРМАЦІЙНОГО ПОШУКУ

Семантичний механізм став результатом довгих науково-дослідних робіт (протягом 8 років) в області пошукової лінгвістики. У цей процес було вкладено значні кошти (8-мільйонів \$). Усе почалося з розробки класифікаційної системи, призначеної для даних, зібраних для першого видання Кембриджської енциклопедії. Пізніше система одержала широке поширення серед багатьох інших енциклопедій різних видавництв.

Семантичний пошук – процес пошуку документів за їхнім змістом.

Для здійснення семантичного пошуку необхідні наступні умови: 1) переклад змісту документів і запитів із розмовної мови на інформаційно-пошукову мову й складання пошукових образів документа й запиту; 2) складання пошукового опису, в якому вказується додаткова умова пошуку.

Семантичний пошук розподіляє сторінки не тільки за наявністю в них слів і фраз, які містяться в пошуковому запиті, - він також «розуміє» зміст сторінок і може видавати за запитом, наприклад, сторінки, що містять синоніми слова-запиту.

Всі технології семантичного пошуку дозволяють користувачам задавати довільні складні питання, а потім інтерпретують їх і застосовують до наявних баз даних. Такими базами є Nakia, Powerset, Freebase тощо, і всі вони мають систему автоматичної обробки розмовної мови, яка "переводить" питання на стандартний запит, зрозумілий для бази.

На сьогоднішній день існують лише кілька областей знань, де семантичний пошук дійсно справляється найкраще – це складні запити про висновки й міркування про складні системи даних.

Для того, щоб технології семантичного пошуку зайняли свою нішу на ринку, розроблювачам необхідно запропонувати зовсім новий інтерфейс, щоб користувачі змогли повністю відчути всю міць семантичного пошуку.

UDC 681.3

G.S. Kralina, E.V. Tupota
National Aviation University

THE INFORMATION RETRIEVAL TECHNOLOGY

Information retrieval is the science of searching for documents, for information within documents and for metadata about documents, as well as that of searching relational databases and the World Wide Web. There is overlap in the usage of the terms data retrieval, document retrieval, information retrieval, and text retrieval, but each also has its own body of literature, theory, praxis and technologies. IR is interdisciplinary, based on computer science, mathematics, library science, information science, information architecture, cognitive psychology, linguistics, statistics and physics. Automated information retrieval systems are used to reduce what has been called "information overload". Many universities and public libraries use IR systems to provide access to books, journals and other documents. Web search engines are the most visible IR applications.

An information retrieval process begins when a user enters a query into the system. Queries are formal statements of information needs, for example search strings in web search engines. In information retrieval a query does not uniquely identify a single object in the collection. Instead, several objects may match the query, perhaps with different degrees of relevancy. An object is an entity which keeps or stores information in a database. User queries are matched to objects stored in the database. Depending on the application the data objects may be, for example, text documents, images or videos. Often the documents themselves are not kept or stored directly in the IR system, but are instead represented in the system by document surrogates.

Most IR systems compute a numeric score on how well each object in the database match the query, and rank the objects according to this value. The top ranking objects are then shown to the user. The process may then be iterated if the user wishes to refine the query. Many different measures for evaluating the performance of information retrieval systems have been proposed. The measures require a collection of documents and a query. All common measures described here assume a ground truth notion of relevancy: every document is known to be either relevant or non-relevant to a particular query. In practice queries may be ill-posed and there may be different shades of relevancy.

УДК 004.424.42(043.2)

Д.С. Газізов, О.І. Курченко

Національний авіаційний університет

ПОШУК ІНФОРМАЦІЇ, РЕЛЕВАНТНОЇ ПОТРЕБАМ КОРИСТУВАЧА

Інформаційно-пошукові системи з'явилися давно. Теорії та практиці побудови таких систем присвячена досить велика кількість статей. Серед вітчизняних джерел слід виділити науково-технічну збірку «Науково-технічна інформація». Російською мовою видана так само і «біблія» по розробці цього роду систем – «Динамічні бібліотечно-інформаційні системи Жерарда Солтона» (Gerard Salton), в якій розглянуті основні принципи побудови інформаційно-пошукових систем і моделювання процесів їх функціонування. Таким чином не можна сказати, що з появою Internet і бурхливим входженням його в практику інформаційного забезпечення, з'явилося щось принципово нове, чого не було раніше.

Головне завдання інформаційно-пошукової системи - це пошук інформації, релевантної інформаційним потребам користувача. Слово релевантність означає відповідність між бажаною і дійсно отриманою інформацією. Релевантність можна ще представити як міру близькості між реально отриманими документами і тим, що слід було б отримати з системи. Природно, що виникає два завдання, які слід вирішити: представлення інформації в системі і формулювання інформаційних потреб користувача.

Найбільш поширеними моделями представлення документів в інформаційно-пошуковій системі є різні варіації на тему векторної моделі, коли документ представляється як набір термінів. Багато систем застосовують механізм корекції запиту по релевантності. Це означає, що процедура пошуку носить інтерактивний і ітеративний характер. Після проведення первинного пошуку користувач відзначає зі всього списку знайдених документів релевантні. На наступні ітерації система розширює/уточнює запит користувача термінами з цих документів і знову виконує пошук. Так продовжується до тих пір, поки користувач не відзначить, що кращого результату, чим він вже має досягти не можна. Корекція запиту по релевантності - це достатньо широкий упроваджений спосіб уточнення запитів.

УДК 519.8:681.5

**Л.И. Гришина, Б.Н. Нестеренко, к.т.н.,
И.А. Шахматов**

Национальный авиационный университет

ПОСТАНОВКА ЗАДАЧИ ДИСПЕТЧЕРСКОГО УПРАВЛЕНИЯ ПАРКОМ МАШИННО- ТЕХНОЛОГИЧЕСКИХ СИСТЕМ

Рассматриваются территориально распределенные в некотором регионе процессы природного характера, которые происходят стихийно или по воле человека и поддаются прогнозированию и/или планированию.

Участие человека в этих процессах происходит в форме применения машинной технологии, представленной совокупностью специализированных агрегатов, самоходных или подключаемых к мобильному энергетическому средству.

Каждый процесс характеризуется некоторым объемом работ. Этот объем работ имеет оптимальный во времени период его реализации с применением машинной технологии. Величина этого объема работ меняется во времени и измерима инструментальными средствами наземного, воздушного и космического базирования.

Процесс выполнения объема работ подвержен воздействию окружающей среды, которая существенно влияет на сроки и качество выполнения запланированного объема работ. Прогноз поведения внешней среды в районах размещения вышеуказанных процессов может быть получен с известной точностью.

Агрегаты составляют некоторый парк, т.е. совокупность агрегатов объединенных единым центром управления. Эти агрегаты размещены в территориально распределенных местах стационарного базирования, причем в каждом месте базирования находится один или несколько агрегатов.

Координаты мест протекания процессов и размещения агрегатов известны с достаточной точностью. Эти места соединены сетью дорог разного качества и разной пропускной способности.

Задача: найти оптимальное динамическое распределение (с учетом перемещения) агрегатов во времени и на множестве мест протекания процессов.

УДК 004.01(082)

М.О. Ментю

Національний авіаційний університет

АКТУАЛЬНІСТЬ ПРОБЛЕМИ ПОШУКУ ІНФОРМАЦІЇ

Сьогодні спостерігається експонентне зростання кількості джерел інформації у світі, що зумовлено збільшенням чисельності її споживачів, обсягу створеної і доступної інформації. Це викликає все більші складнощі в ефективному пошуку інформації, що є наслідком, з одного боку, особливостей людино-машинної взаємодії, а з іншого – семантичної неоднорідності джерел інформації. Розв'язання цієї проблеми полягає в індивідуалізації засобів інформаційного пошуку, тобто в адаптації процесу пошуку до індивідуальних особливостей користувачів, що дозволить швидко знаходити релевантну інформацію час за допомогою прийнятних засобів і з мінімальними зусиллями користувачів.

Значний внесок у розробку теоретичних і прикладних питань підвищення ефективності інформаційного пошуку здійснили Г. Селтон, Дж. Солтон, Е. Ф. Скороходько, Л. Е. Пшенична, В. В. Сидоренко, В. М. Дріянський, О. Г. Дубінський, Ю. В. Рогушина та ін. Однак розроблені методи адаптивного інформаційного пошуку, такі як зворотній зв'язок за релевантністю, різні методи модифікації запитів користувачів, недостатньою мірою враховують специфіку інформаційного пошуку в Інтернеті. Такому пошукові притаманна висока вартість зв'язку, децентралізованість, неоднорідність і різноманітність інформаційних ресурсів, а також небажання користувачів витратити час та зусилля на використання методів покращення інформаційного пошуку. З погляду мінімізації часу і вартості пошуку найбільш перспективним є використання моделей користувачів для розширення запиту під час проведення адаптивного інформаційного пошуку, що дозволяє суттєво скоротити час інтерактивної взаємодії і витрати користувачів, оскільки уточнення запиту виконується на боці пошукового засобу без залучення посередника-людини.

УДК 004.72(043.2)

Е.А. Зубарева

Национальный авиационный университет

МЕТОД ПЕРЕДАЧИ МУЛЬТИМЕДИЙНОГО ТРАФИКА В БЕСПРОВОДНЫХ МУЛЬТИСЕРВИСНЫХ СЕТЯХ

На сегодняшний день не решен вопрос создания сетей, поддерживающих режим видеоконференцсвязи (ВКС) с обеспечением заданных параметров качества обслуживания (QoS) и с требуемым уровнем помехозащищенности. Для реализации режима ВКС необходимо решить две основные задачи:

- синхронизировать мультимедийный трафик (ММТ);
- обеспечить помехозащищенность функционирования сети, поддерживающей передачу ММТ.

Важно отметить, что для решения этих задач необходимо разработать метод реконфигурации диаграмм направленности антенн (ДНА), в котором обеспечивается своевременная синхронная передача ММТ в условиях недостаточной помехозащищенности. Для этого, необходимо установить взаимосвязь двух стохастических процессов: передачи синхронизируемого ММТ при условии отсутствия помех и передачи ММТ с учетом реконфигурации ДНА, обеспечивающей переход в зону радиоспектра без помех.

В процессе передачи ММТ с заданным уровнем QoS должны обеспечиваться жесткие требования к синхронности и безошибочности передачи. Кроме того, в беспроводных сетях первоочередной задачей является повышение помехозащищенности функционирования линий связи (ЛС), за счет реконфигурации ДНА.

В связи с этим, рассмотрено взаимодействие двух процессов: передачи ММТ и коррекции ДНА, обеспечивающей уход из зоны помех. Обратим внимание, что в обоих случаях есть параметры, которые могут не корректироваться, но существуют и такие, которые подлежат обязательной корректировке для обеспечения эффективной передачи ММТ.

В результате можно сделать вывод, что решение двух поставленных задач обеспечит создание метода передачи ММТ в беспроводных мультисервисных сетях за счет применения эффективной реконфигурации ДНА для обеспечения помехозащищенного функционирования..

УДК 658.012.56

О.Є. Литвиненко, д.т.н., А.В. Полухін, к.т.н.,

А.Б. Чапкіс, к.е.н.

Національний авіаційний університет

ПРИНЦИПИ ПОБУДОВИ ТИПОВОЇ КОМП'ЮТЕРИЗОВАНОЇ СИСТЕМИ ПОТОЧНОГО ТА ПІДСУМКОВОГО КОНТРОЛЮ ЗНАТЬ СТУДЕНТІВ

Сучасна концепція проектування інформаційних систем для вищої школи вимагає застосування системного підходу, який розглядає організацію навчального процесу як єдиний процес, що складається з ряду окремих взаємопов'язаних компонентів. Кожен з цих компонентів реалізується у відповідному функціональному блоці системи, що проектується. Функціональний блок в загальному випадку може взаємодіяти з іншими як на рівні процесів, так і на рівні обміну даними. Такий підхід дозволяє здійснювати контроль за навчальним процесом, приймати оперативні рішення з метою підвищення якості підготовки фахівців, накопичувати поточну інформацію про академічну успішність кожного студента та автоматизувати формування підсумкових документів – додатків до диплому.

На підставі аналізу вимог до типової комп'ютеризованої системи поточного та підсумкового контролю знань студентів (ППКЗС) вищого навчального закладу запропонована її загальна функціональна структура. Визначено, що вона повинна складатися з п'яти основних підсистем: підсистеми ведення нормативно-довідкової інформації навчального закладу; підсистеми ведення навчального процесу (модуля ведення організаційної інформації та модуля ведення поточної інформації); підсистеми аналізу знань студента; підсистеми формування звітних документів; підсистеми порівняльного аналізу електронних текстів студентських робіт.

У відповідності до запропонованої функціональної структури системи ППКЗС створено UML-модель, в якій детально спроектовані діаграми варіантів використання як всієї системи в цілому, так і її окремих підсистем. Для кожної з підсистем побудовані діаграми класів. Здійснена програмна реалізація алгоритмів функціонування системи ППКЗС.

УДК 004.724.4(045)

І.А. Жуков, д.т.н.,
І.А. Клименко, к.т.н., В.О. Саченко
Національний авіаційний університет

МЕТОД РЕПЛІКАЦІЇ ДАНИХ В МОБІЛЬНИХ GRID

Проблемам сумісності *Grid*-систем і мобільних технологій та створенню мобільних *Grid*, присвячена велика кількість сучасних досліджень [1]. Динамічна природа мобільної *Grid*, крім необхідності динамічного планування ресурсів, обумовлює також необхідність дублювання даних. Вузол, який обробляє чергову порцію даних, може вийти із складу мережі, що спричинить втрату результатів і помилки обчислювального процесу. Виникає необхідність розробки нових методів реплікації із урахуванням характерних особливостей мобільних *Grid*.

Для дослідження можливих рішень проблеми проведений огляд методів реплікації даних. Розглянуті існуючі методи реплікації «*master-slave*», «*client-server*», «*peer-to-peer*» [2], можливість їх модифікації для мобільних *Grid*. В результаті проведеного аналізу стало очевидно, що існуючі механізми не адаптовані до динамічного обчислювального середовища, що знижує їх надійність та ефективність. Проведені дослідження показали, що найбільш сприятливою для використання в мобільних *Grid* є модель «*peer-to-peer*». Проблемами такої моделі під час її використання в мобільних *Grid* є великий об'єм реплікованих даних та її низька експлуатаційна якість.

Для побудови мобільної мережі і здійснення реплікації даних *Grid* необхідно два алгоритми – підтримки дерева мережі і пошуку місцезнаходження даних. Запропонований алгоритм підтримки дерева. Кожен вузол підтримує список *привілейованих сусідів*. Дерево будується з головного вузла, який підключений до мережі. Нові вузли вибираються із списку доступних вузлів, використовуючи різні метрики. Алгоритм заснований на використанні гібридної моделі даних, найбільш оптимальної для мобільної *Grid* з погляду її топології.

Список літератури

1. Кулаков Ю.О., Клименко І.А. Компоненты и аспекты динамической GRID среды // 36. Наук. Праць «Проблеми інформатизації та управління». – № 3(27). – р. 34 – 41.
2. Lamahemedi H. Decentralized data management framework for data grids. – Rensselaer Polytechnic Institute Troy, New York. – ст. 22–24

В.Є. Снитюк, к.т.н.

Черкаський державний технологічний університет

ЕВОЛЮЦІЙНІ ТЕХНОЛОГІЇ ПРИЙНЯТТЯ РІШЕНЬ В УМОВАХ НЕПОВНОЇ АПРІОРНОЇ ІНФОРМАЦІЇ

Останні роки не були відзначені відомими розробками та відкриттями у сфері технологій прийняття рішень, пов'язаних із використанням природно-орієнтованих механізмів, якими стали нейронні мережі, еволюційне моделювання та методи теорії нечітких множин.

Сучасні задачі ідентифікації і прогнозування, які базуються на основі аналізу таблично заданих даних та розглядаються при проектуванні складних технічних систем, інформаційно-аналітичних систем, різного роду дослідженнях, розв'язувати з використанням класичних методів є проблематичним, оскільки сучасні економічні процеси відрізняються стрімкою динамікою та якісними змінами. Наслідком цього є їх нестаціонарність та об'єктивна і суб'єктивна невизначеність. Об'єктивну невизначеність складають неповнота, неоднозначність даних, які є наслідком впливу технічних та природних факторів. Суб'єктивний характер інформації, її неоднозначність та невстановлена компетентність експертів є елементами суб'єктивної невизначеності.

Практична реалізація такої парадигми базується на розв'язанні задач ієрархічної схеми послідовного подолання невизначеності. Розв'язання задачі відновлення пропусків є необхідною умовою ідентифікації залежності, але за умови, що дані мають однорідний характер. Якщо це не так, то табличні дані потрібно кластеризувати. І оскільки їх структура невідома, то запропоновано еволюційний метод EvoClust, який дозволяє здійснювати кластеризацію за мінімальних початкових умов. Таким чином, в умовах невизначеності задачі ідентифікації доцільно розв'язувати, використовуючи нейромережі як універсальні моделі залежностей, та методи еволюційного моделювання для оптимізації їх параметрів і оптимізації цільових функцій.

Список літератури

1. *Згуровский М.З., Панкратова Н.Д.* Системный анализ. Проблемы, методология, приложения. – К.: Наукова думка, 2005. – 743 с.
2. *Снитюк В.Є.* Прогнозування. Моделі, методи, алгоритми. – К: Маклаут, 2008. – 364 с.

УДК519.2.+600.1

И.И. Горбань

Институт проблем математических машин и систем НАНУ

ФИЗИКО-МАТЕМАТИЧЕСКАЯ ТЕОРИЯ ГИПЕРСЛУЧАЙНЫХ ЯВЛЕНИЙ

Цель доклада – познакомить участников конференции с новой физико-математической теорией гиперслучайных явлений [1].

Главная особенность этой теории – учет статистической нестабильности физических явлений (событий, величин, процессов и полей), наблюдаемых в реальном мире.

Теория гиперслучайных явлений имеет математическую и физическую составляющие. Математическая часть базируется на аксиомах теории вероятностей, а физическая – на новых гипотезах, отличающих ее от классической теории вероятностей:

1) *аксиоме статистической непредсказуемости* явлений, означающей наличие в мире непредсказуемых явлений, происходящих вне связи с ранее происходившими явлениями;

2) *гипотезе о гиперслучайном характере реального мира (гипотезе гиперслучайности)*, означающей, что аксиома статистической непредсказуемости выполняется повсеместно.

Математическая часть теории гиперслучайных явлений разработана для гиперслучайных событий, величин, функций и функционалов. Физическая часть теории охватывает широкий круг вопросов, связанных с описанием реальных явлений с помощью гиперслучайных моделей.

С помощью теории гиперслучайных явлений выяснена причина ограниченной точности любых физических измерений. Установлено, что кроме систематической погрешности измерений, величина которой остается постоянной при проведении ряда измерений, и случайной погрешности, величина которой уменьшается при накоплении данных, существует еще непредсказуемая погрешность. Величина этой погрешности может меняться от измерения к измерению, но уменьшить ее путем накопления невозможно.

Список літератури

1. *Горбань И.И.* Теория гиперслучайных явлений. – Киев, ИПММС НАН Украины, ГП «УкрНИУЦ», 2007. – С. 184.

УДК 004.724.4(045)

В.В. Ткаченко, к.т.н., А.В. Левчук, Д.С. Пономарчук
*Національний технічний університет «КПІ»
Національний авіаційний університет*

ВИРІШЕННЯ ЗАДАЧ УПРАВЛІННЯ З ЗАСТОСУВАННЯМ КЛАСТЕРНИХ ОБЧИСЛЕНЬ

Сучасний рівень розвитку технологій призводить до стрімкого ускладнення задач керування та підвищення вимог до автоматичних систем управління (АСУ). Одним із відомих шляхів підвищення ефективності АСУ є підвищення обчислювання потужності системи за рахунок нарощування кількості процесорів [1]. Модульний принцип організації таких систем відповідає вимогам масштабування апаратного і програмного забезпечення. У такий спосіб необхідність розв'язку задач обумовлює розробку нових обладнань керування або вдосконалення вже існуючих. Використання кластерних обчислень для розв'язання задач управління дозволить організувати обробку управляючої інформації за рахунок використання паралельних обчислень, а також підвищити універсальність АСУ. Адаптація таких систем управління до різноманітних об'єктів управління можлива за рахунок використання високошвидкісних міжсистемних з'єднань та реалізується на рівні ПЗ [2].

Найчастіше задачі управління зводяться до вирішення траєкторних задач. До методів вирішення таких задач відносяться: методи оціночної функції, метод цифрових диференціальних аналізаторів, метод «цифра за цифрою», таблично-алгоритмічний метод [3].

В рамках зазначених досліджень була розроблена програма розв'язку системи лінійних рівнянь методом Якобі мовою C++ з використанням технології MPI. Основна розрахункова формула має вигляд: $x^{(n+1)} = Bx^{(n)} + d$, де основою виразу є перемноження матриці на вектор, що паралельно обчислюється на безлічі обчислювальних вузлів кластера.

Список літератури

1. *Жабін В.І.* Архітектура обчислювальних систем реального часу. – К.: Століття+, 2003. – 176 с.
2. *Perkins C., Gharai L.* Real-time Collaborative Environments and the Grid // Proc. of the Workshop on Advanced Collaborative Environments (WACE'04). – Nice, France, 2004.
3. *Байков В.Д., Вашкевич С.Н.* Вирішення траєкторних задач в мікропроцесорних системах ЧПУ/Л.: Машинобудування, 1986. – 106 с.

УДК 004:371.279.7

Т.В. Шепелева, к.т.н., А.О. Медведчук
Національний авіаційний університет

АРХІТЕКТУРА КОМП'ЮТЕРИЗОВАНОЇ СИСТЕМИ ПОТОЧНОГО ТА ПІДСУМКОВОГО КОНТРОЛЮ ЗНАТЬ СТУДЕНТІВ ВИЩИХ НАВЧАЛЬНИХ ЗАКЛАДІВ

При системному проектуванні програмної системи, формуванні її архітектури, виборі компонент і формалізації зв'язків між її складовими частинами слід враховувати ряд вимог:

– архітектура системи повинна відповідати поточним і перспективним цілям та функціональним стратегічним завданням системи;

– для забезпечення перспективи розвитку системи слід передбачати можливість інтеграції компонент на різні апаратні і операційні платформи на основі концепції і стандартів відкритих систем;

– архітектура програмної системи має бути достатньо гнучкою і допускати відносно простий, без суттєвих структурних змін, розвиток і нарощування функцій і ресурсів системи відповідно до розширення сфер і завдань її застосування;

– забезпечити ефективне використання ресурсів системи і мінімізувати інтегральні витрати на обробку даних в типових режимах її функціонування з урахуванням поточних експлуатаційних витрат;

– повинні бути передбачені безпека функціонування комплексу програм і надійний захист даних від помилок, від навмисного руйнування або втрати інформації, а також авторизація користувачів, управління робочим завантаженням, резервуванням і оперативним відновленням функціонування системи;

– слід забезпечити комфортний, максимально спрощений доступ кінцевих користувачів до управління і результатів функціонування системи на основі сучасних графічних засобів і наочних призначених для користувача інтерфейсів.

З урахуванням цих вимог було прийняте рішення обрати для системи, що проектується, 3-рівневу клієнт-серверну архітектуру: клієнт, сервер та база даних. Розроблено структурну схему системи і проєкт-схема її бази даних. Для реалізації деяких функцій системи розроблені сервіси, що не будуть частиною програмного серверу.

УДК 004.3

В.В. Жабина

Национальный технический университет Украины «КПИ»

ОРГАНИЗАЦИЯ ПАРАЛЛЕЛЬНЫХ ПОТОКОВ КОМАНД В СИСТЕМАХ, УПРАВЛЯЕМЫХ ПОТОКОМ ДАННЫХ

Одним из перспективных подходов, позволяющих устранить ряд недостатков статического планирования параллельных вычислений, является разработка средств динамического распараллеливания вычислений. Такой подход дает возможность достичь большей степени параллелизма, так как позволяет выявить параллельные ветви, которые возникают непосредственно в процессе вычислений и не могут быть определены при статическом планировании.

Для динамического распределения заданий между вычислительными ресурсами системы предложены модели вычислений, управляемых потоком данных. В таких системах распределение работ между вычислителями при одном потоке команд осуществляется автоматически на аппаратном уровне. Однако распараллеливание потоков команд требует статического анализа, что приводит к недостаткам статических методов.

Предлагается метод параллельного формирования команд, который не требует ручного вмешательства в процесс программирования. Подготовка вычислений осуществляется на основе графа, каждой вершине которого соответствует операция, а каждой дуге – операнд. Метод базируется на получении префиксной бесскобочной формы записи вычисляемого результата в виде строки. В полученной строке выделяются все пары имен операндов вместе с именем соответствующего актора (выполняемой операции). В результате прохода выписываются имена всех акторов, которые могут выполняться параллельно во времени. Далее рассматриваются эти тройки объектов в качестве операндов для следующего прохода.

По сравнению с традиционной моделью вычислений предлагаемый подход упрощает процесс подготовки задач. Акторы и данные могут быть сформированы автоматически компилятором на основе графа. Вычислительные модули системы могут выполнять операции, относящиеся к разным задачам, в любой последовательности, что обеспечивает многозадачный режим обработки данных.

УДК 004.3

В.И Жабин, д.т.н., В.В. Макаров, к.т.н.
Национальный авиационный университет

ТАБЛИЧНО-АЛГОРИТМИЧЕСКИЙ МЕТОД ПОЛИНОМИАЛЬНОЙ АППРОКСИМАЦИИ ФУНКЦИЙ

Для систем числового программного управления характерным является необходимость решения в реальном времени траекторных задач, требующих вычисления функциональных зависимостей, которые могут быть аппроксимированы полиномами. Важной задачей является разработка методов быстрого вычисления полиномов.

Проведен анализ эффективности различных методов вычисления полиномов. Наиболее эффективными являются таблично-алгоритмические методы [1], которые сочетают обращение к таблицам и обработку полученной информации.

Показана возможность ускорения вычислений полиномов с использованием предварительной обработки коэффициентов. Показано, что многочлен $P(x) = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_m X^m$ можно преобразовать к виду $P_1(X) = \alpha_0 + (\alpha_1 + X)^2 + \dots + (\alpha_m + X)^{m-1} + F_1(X)$, который позволяет улучшить характеристики вычислителя. Получены и проанализированы формулы для пересчета коэффициентов.

Рассмотрены аспекты реализации метода с использованием современной элементной базы. Для смены воспроизводимой функции достаточно один раз предварительно пересчитать коэффициенты полинома. Схема устройства при этом не претерпевает изменений.

Полином может быть вычислен за время $T = t_c + t_{\Pi} + t_d$, где t_c – время сложения двух чисел в сумматоре, t_{Π} – время обращения к памяти (таблице), t_d – время сложения $(m+2)$ чисел с помощью дерева сумматоров. Предложенный метод вычисления полиномов позволяет по сравнению с наиболее эффективным из рассмотренных известных таблично-алгоритмических методов ускорить воспроизведение функций.

Список літератури

1. *Кнут Д.* Искусство программирования для ЭВМ: В 3-х т., т. 2.— М.: Мир, 1977.— 723 с.

УДК 004.02

О.М Глазок., к.т.н.

Національний авіаційний університет

ЗАСТОСУВАННЯ ДРУГОГО МЕТОДУ ЛЯПУНОВА ДЛЯ РОЗВ'ЯЗАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАІЧНИХ РІВНЯНЬ ВЕЛИКОЇ РОЗМІРНОСТІ

Багато задач, де використовуються математичні моделі систем з розподіленими параметрами, математичні моделі технічних пристроїв, що складаються з великого числа елементів, сполучених локальними зв'язками, а також задачі економіки, біології та інших галузей науки, зводяться до чисельного розв'язання систем лінійних алгебраїчних рівнянь. Сучасні прикладні задачі нерідко породжують системи, у яких кількість рівнянь може складати сотні тисяч чи навіть мільйони, і потреби практики в розв'язанні задач все більшої розмірності зростають. При цьому розмірність задач, а відтак і обсяг даних, що необхідно зберігати та обробити при їх розв'язанні, зростають швидше, ніж обсяг оперативної пам'яті доступних обчислювальних машин, і швидше, ніж їх швидкісні характеристики. В системах з великою кількістю елементів зв'язки між елементами найчастіше є локальними, і, відповідно, матриці систем лінійних алгебраїчних рівнянь, що описують такі системи, найчастіше є розрідженими. У виконаному дослідженні було запропоновано, по-перше, спосіб раціональної організації обчислювального процесу, який дозволяє добитися значного зменшення потреб у оперативній пам'яті для зберігання даних, а також зменшення тривалості обчислень – завдяки тому, що немає потреби виконувати арифметичні дії, результат яких задалегідь відомий. По-друге, на основі другого методу Ляпунова побудовано допоміжну функцію і запропоновано ітераційний метод, при застосуванні якого отримана послідовність числових векторів сходиться до розв'язку заданої системи. По-третє, на випадок, якщо через велику розмірність задачі навіть за умови застосування запропонованої організації обчислень обсяг даних перевищує можливості доступної обчислювальної техніки, пропонується застосування другого «шару» ітерацій, що дозволяє організувати послідовність отримання розв'язків підмножин системи рівнянь і за рахунок застосування другого методу Ляпунова добитися збіжності цього процесу.

УДК 004.272.1

Є.В. Шевцова

Національний авіаційний університет

СПОСІБ ФУНКЦІОНУВАННЯ СИСТЕМИ МОНІТОРИНГУ ТРАФІКУ БЕЗДРОТОВИХ МУЛЬТИСЕРВІСНИХ МЕРЕЖ

Останнім часом на ринку з'явилася велике розмаїття систем управління і моніторингу бездротових мультисервісних мереж (БММ). Існує ряд причин для такого розвитку, в зв'язку з тим, що деякі нові технології настільки складні, що можуть бути використані тільки в комплекті із спеціально розробленими засобами управління. У зв'язку з цим актуальною є задача розробки системи моніторингу трафіку в БММ.

Для цього треба вирішити задачі: аналізу заголовків рівнів та аналізу результатів вимірювань. Дані задачі можна розв'язати за допомогою розробки загальної схеми аналізу мереженого трафіку та способу функціонування системи моніторингу (рис. 1).

Для розробки системи моніторингу важливими є параметри: IP- та MAC-адреси користувача, об'єм вхідного-вихідного трафіку та час і тривалість з'єднання користувача в мережі.



Рис. 1. Схема функціонування системи моніторингу трафіку в БММ

За схемою та зазначеними вище параметрами планується розробка системи моніторингу трафіку БММ у вигляді самостійного програмного продукту.

УДК 004.724.4(045)

Д.В. Лобурь

Національний авіаційний університет

МОНІТОРИНГ ХАРАКТЕРИСТИК СУПУТНИКОВИХ НАВІГАЦІЙНИХ СИСТЕМ

Навігаційний сервіс GNSS, який надається користувачам, полягає у розповсюдженні навігаційних сигналів супутників GPS, ГЛОНАСС, GALILEO, BEIDOU і диференціальної корегувальної інформації, яка формується функціональними доповненнями супутникових навігаційних систем, як широкозонними (WAAS, EGNOS, Skyfix), так і регіональними (SAPOS, CORS, AGNES, СКНЗУ).

Важливо відмітити, що від виду послуг, які надаються користувачам, залежить склад контрольованих параметрів якості навігаційного забезпечення. Для послуг реального часу серед складових аналізу якості послуг першочерговою є SIS-верифікація, яка використовується з метою навігації. Існують й інші характеристики якості навігаційного забезпечення: безперервність; цілісність; точність.

Звернемо особливу увагу на безпосереднє відношення супутникових навігаційних систем до інтелектуальних інформаційних систем (ІС). ІС є комплексом програмних, лінгвістичних і логіко-математичних засобів для реалізації основного завдання: 1) здійснення підтримки діяльності людини; 2) пошуку інформації в режимі розширеного діалогу на природній мові.

Наведемо основні завдання, які вирішують ІС, серед яких моніторинг – одне з найважливіших завдань.

Основне завдання моніторингу – безперервна інтерпретація даних в реальному масштабі часу і сигналізація про вихід тих або інших параметрів за допустимі межі.

Головні проблеми – «пропуск» тривожної ситуації і інверсне завдання «помилкового» спрацьовування. Складність цих проблем в розмитості симптомів тривожних ситуацій і необхідність обліку тимчасового контексту.

З вищевказаної інформації чітко стає зрозумілим доцільність вивчення та удосконалення супутникових інтелектуальних навігаційних систем з метою їх подальшого впровадження у розвиток науки і технології.

УДК 004.724.4(045)

С.С. Донченко

Национальный технический университет Украины «КПИ»

ОБЕСПЕЧЕНИЕ ЗАДАНЫХ ПАРАМЕТРОВ КАЧЕСТВА ОБСЛУЖИВАНИЯ ДЛЯ ТРАФИКА РЕАЛЬНОГО ВРЕМЕНИ

Для современных компьютерных сетей характерным является наличие трафика разного типа. Каждый вид трафика в зависимости от допустимого уровня задержек предъявляет свои требования к качеству обслуживания *QoS* (*Quality of Service*). Кроме того, важно отметить, что для каждого сетевого приложения существует допустимое значение временной задержки при передаче информации.

Как известно, для приёма и передачи пакетов на маршрутизаторах обычно используется метод *FIFO* (*First In First Out*). Для повышения эффективности обеспечения требуемого уровня *QoS*, будем использовать «умную» очередь, в которой приоритет у пакетов зависит не только от порядка поступления в очередь, а также от типа сервиса, который указан в значении поля *ToS* (*Type of Service*) IP-пакета.

В связи с этим, наиболее эффективно дать высокий приоритет трафику реального времени (*VoIP* пакетам) и низкий приоритет трафику данных и транзакций (пакетам *FTP*, *SMTP* и клиентам файлообменной сети).

В связи с тем, что сетевые ресурсы всегда в дефиците и получить неограниченную пропускную способность невозможно, предлагается улучшить качество входящего сигнала в *VoIP*. Это можно осуществить за счет использования разделения трафика на исходящий и нисходящий потоки. *QoS* нисходящего (*Downstream QoS*) потока позволяет, в отличие от восходящего *QoS* установить приоритет входящего трафика.

В случае высокой нагрузки на канал (например, передаются данные большого размера) нисходящий *QoS* будет динамически ограничивать нисходящий трафик данных, для того чтобы освободить полосу для соединений *Voice over IP*.

Таким образом, можно достичь максимально эффективного способа передачи данных при большой нагрузке на канал и при этом время задержки будет минимальным.

УДК 519.872

О.М. Дишлюк

Національний авіаційний університет

МАТЕМАТИЧНА МОДЕЛЬ РОЗВ'ЯЗАННЯ ЗАДАЧІ АНАЛІЗУ РОБОТИ СОТОВИХ СИСТЕМ ЗВ'ЯЗКУ CDMA-ТЕХНОЛОГІЇ

При організації сотових телефонних мереж доводиться вирішувати проблему доступу до базових станцій. Число користувачів сотових мереж мобільного зв'язку з кожним роком зростає, а отже, виникає задача про підвищення якості передачі інформації.

Сьогодні найпоширенішим методом передачі даних в безпроводових мережах є множинний доступ з кодовим розділенням – CDMA. Кожен вузол мережі посилає дані до розподіленої середі в ті моменти, коли це йому потрібно. Канали трафіку при такому способі розділення середі формуються присвоєнням кожному користувачу окремого числового коду, який розповсюджується по всій ширині частотної смуги. Такий спосіб передачі інформації використовується в сотовому зв'язку та в супутниковій навігації (GPS).

Задача аналізу роботи сотових мереж зв'язку, які функціонують за методом CDMA, відноситься до задач теорії масового обслуговування. Загальною особливістю таких задач є випадковий характер досліджуваних явищ (сеанси зв'язку), оскільки кількість заявок на обслуговування, тимчасові інтервали між їх надходженнями та тривалість обслуговування випадкові.

В якості математичної моделі, яка описує процес функціонування мереж сотового зв'язку на основі методу CDMA, а також аналіз її роботи, застосовуємо систему масового обслуговування з вхідним потоком заявок складної структури. Для таких заявок характерне імпульсне обслуговування, тобто з окремою заявкою виникає випадкова множина імпульсів, кожен з яких, в свою чергу, вимагає неперервного обслуговування на каналі. Кількість імпульсів може бути детермінованою або випадковою. Характерною особливістю систем з такими заявками є те, що обслуговування різних заявок складної структури йдеться в довільному порядку, в результаті чого можуть виникати їх перекриття, що приводить до інформаційних втрат. Для опису таких явищ розроблено алгоритм статистичної реалізації випадкової множини імпульсів, одним із результатів якого є автоматизація розрахунку «інтервалу перетину» двох множин.

УДК 004.724.4(045)

И.И. Давиденко, к.т.н.

Национальный авиационный университет

СПОСОБ ОРГАНИЗАЦИИ ДЕЦЕНТРАЛИЗОВАННОЙ МОДЕЛИ УПРАВЛЕНИЯ ДАННЫМИ В GRID-СИСТЕМАХ

Большинство существующих систем управления *GRID* архитектурой имеют централизованное управление со сложным процессом настройки и обслуживания. В связи с этим, актуальной является проблема разработки новых децентрализованных подходов к решению задачи предоставления доступа к ресурсам *GRID*-системы для групп удаленных пользователей.

Предлагаемый способ организации децентрализованной модели управления *GRID*-системами использует метод доступа к данным основанный на концепции *P2P*-систем, однако не гарантирует надежную доставку информации. Основным преимуществом предложенного способа можно считать простоту добавления и удаления узлов системы, а также отсутствие сложных алгоритмов управления и мониторинга *GRID*-системы.

Однако, чтобы гарантировать доступ к данным с требованиями, которые предъявляются к совместным вычислениям, необходима более простая структура. В качестве такой структуры может быть принята кольцевая топология. Она может быть создана и инициализирована в процессе установления соединения с несколькими соседями, имеющими наибольшую степень связности. Такую топологию можно представить в виде управляющего оверлейного подграфа графа связности всей *GRID*-системы. Этот подграф определяет пути коммуникаций, которые могут использоваться в процессе передачи по сети местонахождения данных. Важно отметить, что в таком подграфе каждый узел имеет связь с множеством других узлов. Используя структурированные оверлейные подграфы можно сформировать основу подхода к построению масштабируемой системы управления и созданию копий для *GRID*-среды.

Предлагается использовать комбинацию топологий деревьев и колец, для построения оверлейного подграфа и маршрутизации запросов доступа для определения местонахождения данных в *GRID*-системе. Эти структуры могут легко приспособиваться к различной динамической топологии. Мгновенно связывающиеся деревья показали достаточную надежность в *P2P*-системах.

УДК 681.3

Г.С. Краліна

Національний авіаційний університет

АНАЛІЗ МЕТОДІВ ТА МОДЕЛЕЙ ПРИЙНЯТТЯ РІШЕНЬ В СИСТЕМАХ ПЛАНУВАННЯ ТА СИТУАЦІЙНОГО УПРАВЛІННЯ

Управління технічним обслуговуванням літальних апаратів корабельного базування припускає реалізацію функцій обробки даних і прийняття рішень у двох часових контурах: контурі планування і контурі оперативного управління.

У першому з них здійснюється календарне планування операцій, що утворюють у сукупності керований технологічний процес. Задача планування ТО ЛАК є різновидом загальної задачі теорії розкладів. Точних методів рішення задач даного класу в даний час не існує. Тому на практиці звичайно використовуються різноманітні евристичні алгоритми, що не володіють властивістю повноти і, відповідно, не гарантують знаходження припустимих рішень задачі у всіх ситуаціях, коли вони об'єктивно існують. Ця обставина значно знижує практичну цінність відомих розробок і вимагає створення нових методів оперативного планування ТО ЛАК, які повинні базуватися на використанні строгих математичних моделей і оптимізаційних алгоритмів, які володіють властивістю повноти.

Функція другого часового контуру реалізується в збійних станах керованого процесу, коли своєчасне завершення всіх запланованих комплексів робіт з ТО ЛАК виявляється неможливим через різні впливи стохастичного характеру. У таких випадках виникає необхідність у відшуканні оперативного управлінського рішення, реалізація якого здатна ліквідувати або звести до мінімуму затримку термінів завершення комплексів взаємозалежних робіт.

Проведений аналіз математичних методів теорії прийняття рішення, показав, що вони носять аксіоматичний і евристичний характер, залишаючи вироблення остаточного рішення за особою, що приймає рішення.

Таким чином, сукупність сучасних інформаційних технологій, дозволяє вести мову про розробку методу прийняття рішень в системах планування та ситуаційного управління, головним призначенням якого є своєчасне і якісне забезпечення всіх інформаційних потреб керівників у процесі ухвалення рішення.

УДК 004.724.4(045)

К.М. Черняк

Національний технічний університет України «КПІ»

ПОХИБКИ ВИМІРЮВАНЬ КООРДИНАТНО-ВИМІРЮВАЛЬНИХ СИСТЕМ

Похибка вимірювання у координатно-вимірювальній системі (КВМ) складається з: похибки збору інформації; похибка вимірювання координат точок; похибки обробки та представлення результатів вимірювань.

Ці похибки суттєво залежать від компонентів, які визначаються складом апаратурної частини, системи програмно-математичного забезпечення координатно-вимірювальної машини, параметрів вимірювальних деталей, факторів навколишнього середовища, джерел енергії, часу вимірювання та експлуатації координатно-вимірювальної машини.

Також треба сказати, що за ступенем впливу на похибку координатно-вимірювальної машини, основними та практично рівноцінними компонентами є механічна конструкція, яка реалізує координатну систему координатно-вимірювальної машини та вимірювальні системи.

В основному, похибка вимірювання залежить від способу вимірювання даних. Зауважимо, що на точність вимірювання суттєво впливає стратегія вимірювання, яка включає: послідовність вимірювання від окремих геометричних параметрів та точок на вимірювальній поверхні, їх кількість та положення на поверхні.

Більшість інноваційних методів вимірювання координатних параметрів можна поділити на два основні способи вимірювання: безконтактний та тактильний.

Безконтактний спосіб ґрунтується на вимірюванні різного хвильового випромінення. Для його вимірювання використовують різні датчики, які сприймають випромінення в певних діапазонах. Тактильний спосіб полягає в визначенні геометричних параметрів певної деталі, через безпосередній контакт робочого органа координатно-вимірювальної машини з предметом дослідження.

Тактильний метод вимірювання може бути застосований при дослідженні відносно малих за розмірами поверхонь.

УДК 004.2 (043.2)

Н.А. Ковалев

Национальный авиационный университет

ОДИН ПОДХОД К ЛОГИЧЕСКОЙ ОБРАБОТКЕ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ОБРАЗОВ

Для решения задач лингвистического анализа и распознавания образов перспективным представляется развитие новых направлений в исследовании булевых функций (БФ), одним из которых выступает позиционная алгебра логики (ПАЛ). Аппарат ПАЛ позволяет компактно представлять БФ при помощи эквивалентных преобразований и позиционных операторов, а также арифметизировать процесс логических вычислений.

Предлагается еще один подход к выражению БФ в рамках ПАЛ. Он заключается в покрытии БФ фрагментами, задаваемыми своими конъюнкциями и описываемыми простыми позиционными операторами с необходимыми коррекциями. Дизъюнктивное объединение подобных записей, минимизация и преобразование полученного выражения к операторной форме приводит к представлению БФ в терминах ПАЛ. При этом разработанный подход характеризуется следующим: 1) использование позиционных операторов невысокой сложности и эквивалентных преобразований только одного вида; 2) существенно меньшая трудоемкость при сохранении компактности представления БФ; 3) возможность совмещения выполнения эквивалентных преобразований и операторов, что сокращает время вычисления БФ.

В соответствии с предложенным подходом для вычисления БФ разработана архитектура комбинационных схем (КС). В сравнении с известными способами реализации БФ при большом числе переменных ($n > 7$) они отличаются значительно меньшей аппаратной сложностью, а в интегральном исполнении – меньшей ресурсоемкостью и большим быстродействием. Масштабируемость архитектуры подобных КС позволяет в процессе работы изменять соотношение между функциональными возможностями и технико-экономическими параметрами микросхем, на базе которых они реализованы, повышая их надежность.

Предлагаемый подход позволяет создавать эффективные средства автоматизации процессов представления БФ от большого числа переменных, синтеза КС и реализации БФ на их основе.

УДК 519.6

В.Ф. Гігіняк

Інститут кібернетики ім. В.М. Глушкова НАН України

ІНТЕРАКТИВНЕ РОЗВ'ЯЗАННЯ ЗАДАЧ ОПТИМАЛЬНОГО КЕРУВАННЯ НЕЛІНІЙНИМИ ОБ'ЄКТАМИ З ФАЗОВИМИ ОБМЕЖЕННЯМИ

Важливим елементом систем підтримки прийняття рішень в авіації, космонавтиці, енергетиці, екології, медицині тощо є інтерактивна підсистема багатоваріантного пошуку оптимальних рішень в задачах керування об'єктами (процесами) різної природи, на функціонування яких накладено додаткові різноманітні обмеження (фазові обмеження). Чим повніше вдається врахувати при постановці і вирішенні задач керування такі обмеження, тим вище якість отриманих рішень.

Розроблена підсистема реалізує одностайний підхід до чисельного розв'язання в інтерактивному режимі задач керування об'єктами різної природи і, відповідно, різних класів математичних моделей. При цьому використовується формалізація фазових обмежень як у метриці L_2 , так і у рівномірній метриці. Це дозволяє включати в математичну постановку задачі умови-обмеження у такій формі, яка найбільш повно і безпосередньо враховує всю різноманітну наявну інформацію щодо реальних процесів.

Розроблений інструментарій було застосовано до розв'язання наступних задач:

- оптимальне керування обертовим рухом твердого тіла при наявності фазових обмежень;
- ідентифікація параметрів математичної моделі розповсюдження ВІЛ-інфекції в Україні;
- відновлення за експериментальними даними граничних умов математичних моделей процесів теплопровідності та дифузії;
- відновлення дійсних сигналів спектроскопії.

УДК 656.71:551.58(082): 681.518.3.09

Н.Ф. Халимон, к.т.н.

Национальный авиационный университет

МЕТОДИКА ОЦЕНКИ КОРРЕЛЯЦИИ И РЕГРЕССИИ МЕТЕОРОЛОГИЧЕСКИХ ЭЛЕМЕНТОВ ПО РЕЗУЛЬТАТАМ МНОГОЛЕТНИХ НАБЛЮДЕНИЙ

Измерения метеорологических параметров на аэродромах являются одним из важнейших элементов системы метеообеспечения взлета и посадки воздушных судов и обеспечения безопасности полетов. В соответствии с этим предъявляются повышенные требования к объему, оперативности и достоверности измерительной метеоинформации и к используемым техническим средствам.

Необходимо проводить регулярные исследования изменчивости высоты нижней границы облаков (НГО), горизонтальной дальности видимости (ГДВ), условий погоды различной степени сложности, опасных явлений погоды, ветра и др. С учетом сравнительно медленного изменения метеоэлементов предложено проводить пересчет климатических характеристик аэродромов.

Для решения задач автоматизации измерений, прогнозирования изменений одних метеоэлементов по результатам анализа других необходимо разрабатывать программные комплексы обработки статистических данных, анализировать характеристики стохастической взаимосвязи наиболее значимых метеоэлементов.

В работе исследованы смешанные центральные моменты высоты нижней границы облаков и горизонтальной дальности видимости в районе аэродрома, находящегося в зоне умеренно-континентального климата. В качестве исходного статистического материала использованы данные многолетних наблюдений и построенная на их основе климатическая характеристика аэродрома. Наиболее гибким и точным методом является аппроксимация полиномами по минимуму среднего квадрата ошибки. В процессе выполнения расчетов установлено, что аппроксимация полиномами четного порядка дает более высокую точность, чем аппроксимация полиномами нечетного порядка, близкого к соответствующему четному порядку. Вообще, порядок полинома, при котором достигается приемлемая точность, сильно зависит от гладкости аппроксимируемой кривой. Поэтому помимо расчета аппроксимирующих функций необходимо выполнять статистическую обработку результатов измерений.

О.М. Бердник

Національний авіаційний університет

МЕТОДОЛОГІЧНІ АСПЕКТИ КОМП'ЮТЕРНОГО МОДЕЛЮВАННЯ АЕРОГІДРОДИНАМІЧНИХ ПРОЦЕСІВ

В роботі розглядається методологічний підхід до комп'ютерного моделювання складних аерогідродинамічних процесів. На сьогоднішній день це один із найрозповсюджених способів наукового дослідження.

Реалізацію завдань запропоновано здійснювати в кілька етапів. Першим етапом є формалізація досліджуваного фізичного явища, тобто побудова математичної моделі. Математичною моделлю аерогідродинамічного процесу є деяка система диференціальних рівнянь з певними значеннями параметрів та певними граничними умовами. В більшості випадків аналітичного розв'язку цих рівнянь знайти не можливо. Спроба ж обмежитися розглядом найпростіших систем, для яких рішення може бути знайдене елементарними методами, суттєво обмежує наші уявлення про фізичний процес.

Другим етапом є побудова комп'ютерної моделі досліджуваного явища, під якою розумітимемо сукупність чисельних методів розв'язку основних рівнянь, алгоритмів їх реалізації та комп'ютерних програм. При цьому результативність комп'ютерної моделі визначається якістю програмного забезпечення, яке буде використовуватись. На нашу думку, потужна спеціалізована система MatLab повністю відповідає висунутим вимогам: простота вводу та корекції вхідних даних, візуалізація (наочність) результатів розрахунку. Зауважимо, що програвання поведінки об'єкту шляхом варіювання деяких параметрів моделі вказує на її імітаційну характеристику. Проте при оцінці результатів обчислювальних експериментів для повноправного їх використання необхідно перевірити адекватність математичних моделей. Пропонуємо здійснювати цей крок шляхом переходу до вже відомих наукових результатів

УДК 656.7.052.001.76

М.В. Лупандин, к.т.н., И.А. Бескровная, И.М. Юрьева
Национальный авиационный университет

МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ПЛАНИРОВАНИЯ ИСПОЛЬЗОВАНИЯ ВОЗДУШНОГО ПРОСТРАНСТВА

В настоящее время проблема безопасности полетов учитывается при анализе использования воздушного пространства (ВП) любого государства. Задача обеспечения безопасности решалась бы просто и однозначно, если бы процессы были детерминированными. На самом деле, как на весь процесс в целом, так и на воздушные суда в частности оказывают воздействие множество случайных факторов.

В докладе рассматриваются особенности моделирования процессов планирования использования воздушного пространства Республики Беларусь в рамках разрабатываемого прикладного программного обеспечения (ППО) автоматизированной системы аэронавигационного обслуживания (АС АНО) Государственного предприятия «Белаэронавигация», предназначенного для автоматизации информационных процессов планирования и контроля использования ВП.

ППО должно обеспечивать отображение на электронном эталоне географической карты составляющих информационной базы АС АНО; моделирование воздушной обстановки; картографический анализ планов полетов.

Моделирование планируемой воздушной обстановки в заданное время или период времени позволит в реальном масштабе времени прогнозировать превышение плановой загрузки секторов над пропускной способностью с дальнейшим формированием предупреждения о превышения интенсивности.

При моделировании на всех этапах производится анализ конфликтных ситуаций (взаимное расположение объектов по времени, месту и высоте). При возникновении конфликта обеспечивается цветовое указание на карте мест возникновения конфликтных ситуаций с указанием конфликтующих объектов и их параметров с помощью информационных формуляров.

Моделирование процессов планирования использования ВП используется для мониторинга безопасности воздушного движения, то есть для выявления «узких» мест и последующего формирования эффективных мероприятий по их устранению.

УДК 004.75

А.Н. Березовский

Национальный авиационный университет

МОДЕЛИРОВАНИЕ ВЫЧИСЛЕНИЙ В КЛАСТЕРНЫХ СИСТЕМАХ

Моделирование, как инструмент исследований сегодня используется практически во всех отраслях науки. Результаты моделирования, такие как точность и время прогноза, определяются дискретностью или размерностью исходной модели.

Для большинства систем именно высокая размерность или вычислительная сложность модели является существенным ограничением при реализации практических расчетов.

Параллельные алгоритмы, должны обязательно учитывать то, что использование межзловых обменов существенно снижает эффективность использования вычислительного пространства.

Современный подход к организации вычислений в распределенных системах основывается на использовании моделей вычислений, которые отображают архитектуру системы и на использовании моделей программирования, которые позволяют эффективно реализовать модель вычислений [1]. В качестве основы для формирования моделей вычислений предложено использование теории взаимодействующих последовательных процессов (CSP теории). При этом не рассматриваются особенности разработки моделей для масштабируемых кластерных систем.

Предложен и обоснован подход к формированию моделей вычислений, который позволяет описать архитектуру МКСМА, а также описать организацию динамического создания процессов, проверку корректности их поведения и взаимодействия между собой.

Предложенные математические модели вычислений могут быть использованы при разработке распределенных приложений для МКСМА, что позволит повысить качество программных продуктов и сократить время их проектирования.

Список літератури

1. *Топорков В.В.* Модели распределенных вычислений. – М.: ФИЗМАТЛИТ, 2004. – 320 с.

С.М. Станко

Національний авіаційний університет

ВИКОРИСТАННЯ ІНТЕРВАЛЬНИХ МЕТОДІВ ДЛЯ ОПТИМІЗАЦІЇ ІНВЕСТИЦІЙНОГО ПЛАНУВАННЯ

Інвестиційне планування полягає в розробці схеми фінансової діяльності підприємства, організації чи установи. Результатом має стати план заходів щодо платіжних потоків протягом певного часу – терміну планування. Вони складаються з витрат, пов'язаних з інвестиційними зобов'язаннями, та очікуваних прибутків від реалізації даного інвестиційного пакету проєктів. Обсяг цих грошових потоків можливо передбачити лише з певною вірогідністю.

Критеріями оптимізації інвестиційного планування виступають показники чистого приведеного прибутку NPV та внутрішньої норми прибутковості IRR.

Вхідні дані інвестиційного планування являють собою нечіткі множини. Для подальшого розрахунку вони розбиваються на складові α -рівні, представлені чіткими інтервалами. Далі всі операції ведуться засобами інтервальної алгебри.

Вирішення задачі полягає в розрахунку песимістичного та оптимістичного варіантів набору показників. Різниця між відповідними песимістичним і оптимістичним значеннями є найвірогіднішим для даного показника.

Інтервальне вираження NPV та IRR є більш інформативним для прийняття рішення, ніж їхні чіткі значення. Воно дозволяє аналізувати багатокритеріальність цих показників.

Застосування інтервального підходу для розв'язання задач з інвестиційного планування, обумовлених невизначеністю параметрів, є більш результативнішим у порівнянні зі звичайним ймовірнісним підходом. Безпосереднє опрацювання нечітких даних засобами інтервальної математики дозволяє розв'язати задачі, які неможливо вирішити засобами ймовірнісно-статистичної теорії.

ДЛЯ ПОДАТК

Наукове видання

**МІЖНАРОДНА НАУКОВО-ТЕХНІЧНА
КОНФЕРЕНЦІЯ
«ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ
ЛІНГВІСТИЧНОГО АНАЛІЗУ»**

21-22 жовтня 2009 року

Тези доповідей

*Тези надруковані в авторській редакції однією із трьох робочих мов
конференції: українською, російською, англійською*

Підп. до друку 10.11.09 Формат 60x84/16. Папір офс.
Офс. друк. Ум. друк. арк. 4,42. Обл.-вид. арк. 4,75
Тираж 100 прим. Замовлення №

Видавництво Національного авіаційного університету «НАУ-друк»
03680. Київ-58, проспект Космонавта Комарова, 1

Свідоцтво про внесення до Державного реєстру ДК №977 від 05.07.2002