

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

Москвін Олексій Михайлович

УДК 681.5:519.876.2

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ МУЛЬТИАГЕНТНОЇ ОПТИМІЗАЦІЇ ІНТЕРНЕТ-
РЕСУРСІВ**

05.13.06 – інформаційні технології

Автореферат

дисертації на здобуття наукового ступеня
кандидата технічних наук

Вінниця - 2013

Дисертацією є рукопис.

Роботу виконано у Вінницькому національному технічному університеті Міністерство освіти і науки України.

Науковий керівник:

доктор технічних наук, професор
Дубовой Володимир Михайлович,
Вінницький національний технічний університет,
завідувач кафедри комп'ютерних систем управління

Офіційні опоненти:

доктор технічних наук, професор
Харченко Вячеслав Сергійович,
Національний аерокосмічний університет
ім. М. Є. Жуковського "Харківський авіаційний інститут"
завідувач кафедри комп'ютерних систем та мереж

доктор технічних наук, професор
Казимир Володимир Вікторович,
Чернігівський державний технологічний університет,
завідувач кафедри інформаційних і комп'ютерних систем

Захист відбудеться “20” 04 2013 р. о 12:00 годині на засіданні спеціалізованої вченої ради Д 05.052.01 у Вінницькому національному технічному університеті за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95, ауд. 210, ГНК.

З дисертацією можна ознайомитись у бібліотеці Вінницького національного технічного університету за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95, ГНК.

Автореферат розісланий “19” 03 2013 р.

Учений секретар
спеціалізованої вченої ради

С. М. Захарченко

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Завдяки масовій комп'ютеризації та поширенню нових інформаційних технологій Інтернет є одним з найважливіших джерел інформації. Мережа веб характеризується надвеликою кількістю ресурсів, які знаходяться у стані постійних структурних та інформаційних змін. Особливістю даного процесу, з точки зору організації матеріалу є те, що ці зміни хаотичні, а структура не адаптована для перегляду користувачами, що призводить до збільшення кількості маршрутів та їх довжини, якими необхідно пройти для знаходження тієї чи іншої інформації, що відповідає меті пошуку. Для вирішення даної проблеми необхідна розробка засобів оцінювання, критеріїв оптимальності, методів і технологій оптимізації гіпертекстової структури веб-ресурсів. Для цього, в основному, використовуються такі підходи: перший – пов'язаний з використанням семантичної інформації та засобів концепції Semantic Web, другий – передбачає використання гіпертекстової метрики для оцінювання та порівняння навігаційних характеристик гіпертекстових структур, третій підхід базується на використанні систем, що надають статистичну інформацію про діяльність користувачів у гіпертекстовій структурі певного Інтернет-ресурсу.

Перевагою першого підходу, що передбачає використання концепції семантичної мережі Semantic Web, значний внесок в розвиток якого зробили Bernes-Lee T., Hendler J., Lassila O., є використання метаданих, що однозначно характеризують властивості і зміст документів замість використовуваного в даний час синтаксичного аналізу, що дозволяє здійснювати аналіз коректності існуючих гіпертекстових посилань з точки зору їх предметної області. Також слід відмітити наявність розвиненого програмного інструментарію, що, однак, не містить засобів роботи з метаданими, створених різними авторами без використання сторонніх засобів їх попередньої інтеграції, що обумовлює складність автоматизованих засобів їх обробки та аналізу.

В основі другого підходу, запропонованого в роботах Botafogo R. A., Rivlin E., Shneiderman B., Кедрова Г. Е., лежить використання гіпертекстової метрики, за допомогою якої можна оцінити гіпертекстову систему на основі формальних показників. Однак дана метрика характеризується відсутністю критеріїв оптимальності гіпертекстової структури, тобто дає змогу оцінити структурні особливості того чи іншого веб-ресурсу, здійснити їх порівняння, але не дає уявлення про навігаційну якість гіпертекстової структури.

Автоматизовані аналітичні системи, побудовані на основі третього підходу, прикладом яких є Google Analytics, дозволяють встановити цільову аудиторію веб-ресурсу, сторінки, що зацікавили користувачів найбільше, та зв'язки з суміжними веб-ресурсами, що є основними джерелами у формуванні аудиторії. Недоліком таких систем є відсутність дорадчих компонентів, які б пропонували користувачеві альтернативні шляхи перегляду, а власнику ресурсу комплексні рекомендації з модифікації метаданих чи гіпертекстової структури для досягнення ресурсом певних цілей.

Беручи до уваги розглянуті підходи, слід зазначити, що вони лише частково вирішують зазначену проблему, яка не знаходить вирішення на рівні безпосередньо сайтів та структури їхніх гіпертекстових зв'язків. Тобто, зазначені засоби вирішують часткові локальні проблеми пошуку та ідентифікації даних при відсутності зворотного впливу на веб-ресурси. Проблема ускладнюється ще й тим, що існуючі засоби ідентифікації засновані, в основному, на аналізі синтаксичної складової гіпертексту, призначеної для користувача, а не для автоматизованих систем, що суттєво погіршує якість результату.

Підсумовуючи вищесказане, можна зробити висновок, що основною проблемою для вирішення задачі оптимізації гіпертекстової структури веб-ресурсів є відсутність ефективних рішень з аналізу та оцінювання складних гіпертекстових структур, відсутність засобів зворотного впливу на веб-ресурс, відсутність засобів модифікації гіпертекстової структури сайту на основі аналізу семантичної інформації та даних його відвідування користувачами.

Складність і трудомісткість розв'язання даної задачі зумовлюють актуальність пошуку шляхів підвищення ефективності існуючих методів і підходів.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась відповідно до планів науково-дослідних робіт кафедри комп'ютерних систем управління та кафедри моделювання та моніторингу складних систем Вінницького національного технічного університету:

- в рамках договору про творчу співдружність 46/4 “Розробка системи мультиагентної оптимізації Інтернет-ресурсів” (ДР № 0112U002139);
 - в рамках держбюджетної науково-дослідної роботи № 46-Д-269 “Розробка теорії та методів оптимальних рішень в умовах комбінованої невизначеності” (ДР № 0105U002431);
 - в рамках госпдоговірної роботи № 2804 “Побудова та ідентифікація інформаційної моделі сільського зеленого туризму Вінниччини та синтез оптимального web-інтерфейсу для доступу до неї” (ДР № 0106U002996);
 - в рамках госпдоговірної роботи № 2806 “Розробка і впровадження мережної геоінформаційної аналітичної системи комплексного державного моніторингу поверхневих вод Вінницької області” (ДР № 0105U006685);
 - в рамках госпдоговірної роботи № 2812 “Створення Інтернет-порталу для ГІС-систем моніторингу Вінницької області” (ДР № 0107U012437);
 - в рамках держбюджетної науково-дослідної роботи 28-Д-299 “Розробка методів інтеграції математичних моделей природних процесів з геоінформаційними системами природних екосистем” (ДР № 0108U000654)
- у відповідності до пріоритетних напрямків розвитку науки і техніки в Україні.

Мета і задачі дослідження. *Метою роботи* є підвищення швидкості пошуку інформації в мережі Інтернет на основі розробки нової інформаційної технології мультиагентної оптимізації веб-ресурсів.

Для досягнення вказаної мети в роботі розв'язуються такі основні задачі:

- розробка методу оптимізації структури сайтів на основі семантичної інформації;
- розробка моделі гіпертекстового інформаційного середовища, що поєднує структурні характеристики гіпертексту, критерії її оптимальності, семантичну інформацію та дані діяльності користувачів;
- розробка моделі поведінки користувачів у гіпертексті;
- створення мультиагентної технології керування процесом оптимізації структури Інтернет-ресурсів;
- розробка програмного забезпечення для автоматизації процесу мультиагентної оптимізації семантичної структури Інтернет-ресурсів.

Об'єкт дослідження – процеси пошуку інформації в Інтернет.

Предмет дослідження – оптимізація структури Інтернет-ресурсів.

В роботі використані такі **методи дослідження**: теорія графів для моделювання гіпертекстової та семантичної структур, формальний аналіз понять та теорія решіток для формалізації гіпертекстової структури за наявності семантичної інформації, теорія інформації та комбінаторний аналіз для перевірки результатів об'єднання онтологій, методи оптимізації для побудови моделі діяльності користувачів, обчислювальні методи та методи математичного програмування для проведення процесу оптимізації, методи нечіткого аналізу для побудови системи оцінювання якості гіпертекстової структури.

Наукова новизна одержаних результатів. В ході розв'язання поставлених задач були отримані наукові результати.

1. Вперше запропоновано нову модель гіпертекстового середовища, яка поєднує суб'єкти та об'єкти інформаційного пошуку, що дозволило враховувати цілі користувачів при оптимізації гіпертекстової структури веб-ресурсів.

2. Вперше запропоновано метод оптимізації структури сайтів на основі семантичної інформації, який відрізняється функціонуванням в умовах невизначеності відомостей про

структуру мережі в цілому, що дозволяє проводити оптимізацію, базуючись лише на даних про окремі та суміжні до нього ресурси.

3. Вперше запропоновано модель поведінки користувачів у гіпертексті, яка ґрунтується на оцінюванні ентропії веб-системи з множиною користувачів, що дозволило визначити функцію корисності маршруту перегляду інформаційних ресурсів та визначити цілі користувачів.

4. Вперше розроблено мультиагентну інформаційну технологію оптимізації гіпертекстових систем, яка, на відміну від існуючих, використовує інтелектуальних агентів, які для здійснення оптимізації використовують метод оптимізації структури сайтів на основі семантичної інформації, модель гіпертекстового інформаційного середовища і модель поведінки користувачів у гіпертексті.

Практичне значення одержаних результатів роботи полягає у нижчевикладеному:

- розроблено алгоритми ідентифікації цілей користувачів та оптимізації гіпертекстової структури на їх основі;
- розроблено алгоритм оптимізації гіпертекстової структури на основі семантичної інформації;
- розроблено методику мультиагентної оптимізації семантичної структури гіпертекстових систем;
- розроблено алгоритми функціонування інтелектуальних агентів та управління їх координацією для оптимізації гіпертекстової структури;
- розроблено програмне забезпечення “Multiagent hypertext optimization system” для здійснення процесу оптимізації гіпертекстових систем.

Результати дисертаційної роботи впроваджені на підприємствах Kamtech (Норвегія, м. Осло) – інформаційна технологія; ТОВ “Арісент Україна” (Україна, м. Вінниця) – методика мультиагентної оптимізації семантичної структури гіпертекстових систем, алгоритми функціонування інтелектуальних агентів, алгоритми ідентифікації цілей користувачів; Басейнове управління водних ресурсів річки Південний Буг (Україна, м. Вінниця) – алгоритм оптимізації гіпертекстової структури на основі семантичної інформації, алгоритм об’єднання онтологій та методику побудови семантично коректної гіпертекстової структури; у навчальний процес кафедри КСУ Вінницького національного технічного університету – інформаційна технологія.

Особистий внесок здобувача. Всі результати, що складають основний зміст дисертації, отримані здобувачем самостійно. В роботах, опублікованих у співавторстві, здобувачу належать: розробка методу перевірки результатів узгодження онтологічних контекстів гіпертекстових ресурсів за наявності семантичної інформації в [3]; розробка методу оптимізації гіпертекстової структури [2,9]; розробка методу оптимізації гіпертекстової структури на основі семантичної інформації [11,12]; розробка нечіткої системи оцінювання якості гіпертекстових структур для їх оптимізації на основі семантичної інформації [7,10]; розробка алгебри узгодження онтологічних контекстів [1,4,8]; модель поведінки користувача у гіпертексті [5,13]; інтелектуальна мультиагентна технологія оптимізації Інтернет-ресурсів [6,14-16].

Апробація результатів дисертації. Результати дисертаційної роботи доповідались та обговорювались на 14 науково-технічних конференціях: XXXIV-XXXX науково-технічних конференціях професорсько-викладацького складу, співробітників та студентів Вінницького національного технічного університету; V Міжнародній конференції “Інтернет-Освіта-Наука-2006” (ІОН-2006) (м. Вінниця, 2006 р.); в XIII Міжнародній конференції з автоматичного управління “Автоматика-2006” (м. Вінниця, 2006 р.); IX-X Міжнародних конференціях “Контроль і управління в складних системах” (м. Вінниця, 2008, 2010 рр.); XVII Міжнародній конференції з автоматичного управління “Автоматика-2010” (м. Харків, 2010 р.); II-III Міжнародних науково-практичних конференціях “Інформаційні технології та комп’ютерна інженерія” (м. Вінниця, 2010, 2012 рр.); Міжнародній науковій конференції

“Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту (ISDMCI’2012)” (м. Євпаторія, 2012 р.).

Публікації. За темою дисертації опубліковано 16 праць, в тому числі 7 статей надруковано у фахових виданнях (2 статті з яких опубліковано у електронних виданнях), затверджених ВАК України, 7 робіт опубліковано у збірках матеріалів конференцій, отримано 2 свідоцтва про реєстрацію авторського права на твір.

Структура роботи та її обсяг. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел і додатків. Основний зміст викладено на 153 сторінках друкованого тексту, містить 77 рисунків, 10 таблиць. Список використаних джерел містить 151 найменування. Загальний обсяг 190 сторінок.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтована актуальність теми, визначено об’єкт, предмет і методи дисертаційного дослідження, сформульовані мета та задачі роботи, відзначені наукова новизна й практична цінність отриманих результатів.

У **першому розділі** зазначено, що під Інтернет-ресурсами розглядаються веб-сайти та відповідна, асоційована з ними, гіпертекстова інформація. В розділі проведено аналіз та класифікацію існуючих моделей гіпертекстових інформаційних середовищ, засобів моделювання гіпертекстових мереж, аналіз та класифікацію засобів опису семантичної інформації в гіпертекстових мережах і мультиагентних системах. Проаналізовано основні області застосування та недоліки існуючих моделей гіпертекстових інформаційних середовищ, внаслідок чого зроблено висновок, що вони характеризуються відсутністю зв’язку структурної компоненти з семантичною інформацією та користувачами, врахування яких дає можливість перейти від синтаксичного аналізу та аналізу залежностей гіпертекстових посилань до аналізу предметних областей.

На основі проведеного аналізу вибрано напрямок і визначено завдання дослідження.

У **другому розділі** розроблено метод оптимізації Інтернет-ресурсів. Також запропоновано нову модель гіпертекстового інформаційного середовища (1), що враховує суб’єкти та об’єкти інформаційного пошуку у вигляді:

$$\begin{cases} M_{HE} = \langle W, O, U \rangle; \\ W = (V, E); \\ O = \langle C, R, I, A \rangle; \\ U = \langle Y, L, T, J \rangle, \end{cases} \quad (1)$$

де W – гіпертекст, O – онтологія, U – користувачі; V – веб-сторінки, E – зв’язки між ними; O – узгоджена онтологія усіх веб-сайтів, що входять у гіпертекст W ; C – поняття онтології, R – відношення, I – об’єкти, A – аксіоми; Y – множина користувачів, що здійснювали перегляд гіпертексту W , L – множина маршрутів, що здійснили користувачі в процесі пошуку інформації, T – множина значень часу, витраченого на перегляд сторінок V у гіпертексті W кожним користувачем за заданий період спостереження, J – множина точок входу та виходу користувачів гіпертексту W .

Запропонована модель поєднує суб’єкти і об’єкти інформаційного пошуку, що дозволяє враховувати цілі користувачів при оптимізації досяжності інформації.

В роботі запропоновано математичну формалізацію операцій для гіпертекстових ресурсів при наявності семантичної інформації. Результат формалізації подано у вигляді алгебри з восьми операцій – визначення кореня таксономічної ієрархії, визначення поняття, яке породжує таксономічну одиницю, знаходження елементів ієрархії заданого поняття, отримання вектора слівформ для поняття, узгодження онтологій на основі експертних правил, перетворення правил узгодження, узгодження онтологій на основі морфологічного словника, узгодження на основі експертних правил і морфологічного аналізу.

Так, оператор визначення поняття, яке породжує таксономічну одиницю, *instanceOf*

визначено як $instanceOf(n_i) = \{n | n \in V \wedge nR_d n_i\}$, де $i=1, \dots, N$. На рис. 1 зафарбований елемент, що є предком n_5 , відображає результат застосування $instanceOf(n_5)$.

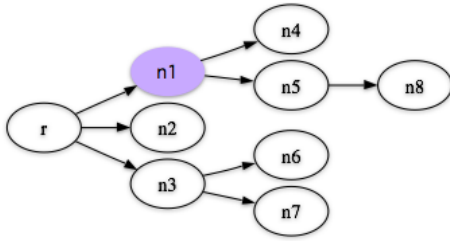


Рисунок 1 – Графічне подання результату застосування оператора $instanceOf(n_i)$

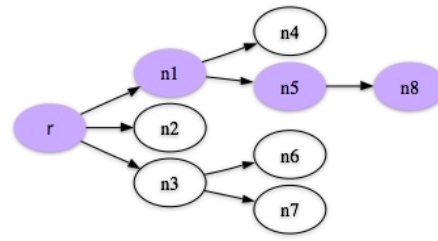


Рисунок 2 – Графічне подання результату застосування оператора χ

Оператор знаходження елементів таксономічної ієрархії заданого поняття $\chi(\alpha, \theta, \lambda)$, де α – поняття, для якого необхідно визначити елементи ієрархії, θ – таксономічна ієрархія онтології O , в якій необхідно здійснювати пошук, λ – рівень поняття в θ . Для визначення даного оператора задамо примітивно-рекурсивну функцію

$$\begin{cases} \chi(\alpha, \theta, 0) = root(\theta); \\ \chi(\alpha, \theta, \lambda) = \alpha; \\ \chi(\alpha, \theta, \lambda - 1) = instanceOf(\chi(\alpha, \theta, \lambda)), \end{cases} \quad (2)$$

де $root$ – оператор відшукування кореня онтології, $instanceOf$ – оператор визначення породного поняття. Результатом застосування оператора $\chi(\alpha, \theta, \lambda)$ є впорядкована множина $\{x | x \in V\}$, де $\inf(\chi(\alpha, \theta, \lambda)) = \alpha$, а $\sup(\chi(\alpha, \theta, \lambda)) = root(\theta)$ – відповідно, нижня та верхня грані множини. На рис. 2 зафарбовані елементи, що формують ієрархію спадкування для n_8 , відображають результат застосування (2) для $\chi(n_8, \theta, 4)$.

Унарний оператор отримання вектора словоформ (3) для поняття онтології $\mu(\alpha(n_i))$ є функціональним відношенням

$$\mu^{-1}(\alpha(V_i)) = \left\{ (m_1, m_2, \dots, m_n) \in U(M) \mid \mu(m_1, m_2, \dots, m_n) = V_i \wedge dom(\mu) = \bigcup_{V_i \in V} \mu^{-1}(V_i) \right\}, \quad (3)$$

де $\alpha(V_i)$ – назва поняття, $U(M)$ – універсальна множина словоформ, (m_1, m_2, \dots, m_n) – множина словоформ для $\alpha(V_i)$.

Оператор узгодження онтологій на основі експертних правил $\omega(O_i, O_j, A_i, A_j)$. A_i – правила узгодження онтологій з суміжних предметних областей, що надаються експертами для здійснення більш точного процесу узгодження і задаються у вигляді множини виразів типу

$(X instanceOf O_1.object) \vee (X has X.option) \vee (O_2 hasConcept Z) \Rightarrow X instanceOf O_2.Z$, де O_1, O_2 – онтології, що узгоджуються, X, Z – поняття, $hasConcept$ – оператор перевірки входження поняття до онтології, has – оператор перевірки наявності певної характеристики у понятті.

Оператор перетворення правил узгодження $\xi(O_1, O_2, A)$ у відношення, що характеризує результат перетворення, тобто $\xi(O_1, O_2, A) \rightarrow (\lambda(O_1.n_j), \lambda(O_2.n_k), \delta) \in R_t$,

де $O_1.n_j$ та $O_2.n_k$, відповідно, поняття з першої та другої онтологій, між якими встановлюються відношення δ , задані у множині правил узгодження A . Функція λ визначає унікальні ідентифікатори понять, які необхідні для коректного узгодження декількох онтологій.

Також визначені оператор узгодження ω на основі правил узгодження та морфологічного словника ω_M . Оператор узгодження на основі експертних правил та

морфологічного аналізу має вигляд

$$\omega_G(O_1, \dots, O_n, A_1, \dots, A_n, M) = O \left(\bigcup_{i=1}^n V_i, \left(\bigcup_{i=1, j=1, i < j}^n \xi(O_i, O_j, A_i \cup A_j) \right) \cup \left(\bigcup_{i=1, j=1, i < j}^n \xi_M(O_i, O_j, M) \right) \cup \left(\bigcup_{i=1}^n R_i \right) \right), \quad (4)$$

де $M = \bigcup_{i=1}^{|V|} \mu^{-1}(\alpha(V_i))$ – множина, що містить словоформи для всіх понять інформаційної

статті, $\xi_M(O_1, O_2, M)$ – оператор перетворення правил узгодження $\xi_M(O_1, O_2, M) \rightarrow (\lambda(O_1.n_j), \lambda(O_2.n_k), isA) \in R_{im}$, isA – відношення еквівалентності, що встановлюється між елементами $O_1.n_j$ та $O_1.n_k$ онтологій O_1 та O_2 .

Запропоновані оператори формують багатоосновну алгебру $G_{ont} = (U, \Omega)$, де комплект задано як $U = \{O_i, A, M\}$, а множину багатосортних операцій як $\Omega = \{root, instanceOf, \chi, \xi, \xi_M, \omega, \omega_M, \omega_G\}$.

Задача оптимізації гіпертекстових систем в роботі розглядається як зміна структури перехресних посилань веб-ресурсів з метою покращення швидкості досяжності контенту користувачами у заданому гіпертекстовому сегменті за умови виконання обмежень на збереження семантичної цілісності існуючої гіпертекстової структури. Для досягнення мети роботи визначено критерій оптимізації, сформований з таких показників гіпертекстової метрики, як індекс інформаційної компактності, індекс стратифікації та частка відсутніх шляхів у вигляді нечіткої бази знань. Хоча за допомогою зазначеної метрики можна здійснити порівняння характеристик навігації веб-ресурсів, її найбільш суттєвими недоліками є неможливість оцінювання якості гіпертекстової структури та здійснення оцінювання її оптимальності, тому використання нечіткої бази знань є виправданим.

Для цього в роботі були проведені дослідження впливу вертикальних, горизонтальних та змішаних зв'язків у межах різних ієрархій та субієрархій гіпертексту на значення критеріїв оцінки гіпертекстової структури та залучені експерти. В результаті узагальнення отриманих результатів була побудована нечітка система оцінювання якості гіпертекстової структури на основі вищезазначених показників.

Для побудови нечіткої системи оцінювання якості використано схему нечіткого висновку Мамдані з трьома лінгвістичними змінними – індексом стратифікації, індексом інформаційної компактності та часткою відсутніх шляхів. Побудована база знань складається з 19 нечітких правил. На рис. 3 зображені поверхні “входи-виходи” для вихідної змінної від комбінацій вхідних змінних. Згідно з отриманими результатами, терму “високий” функції належності вихідної змінної відповідає певна опукла область. Результати проведених досліджень підтвердили, що існує область оптимальних значень компонентів гіпертекстової метрики – $[0.2; 0.85]$ для індексу інформаційної компактності, $[0.1; 0.8]$ для індексу стратифікації та $[0.25; 0.6]$ для частки відсутніх шляхів.

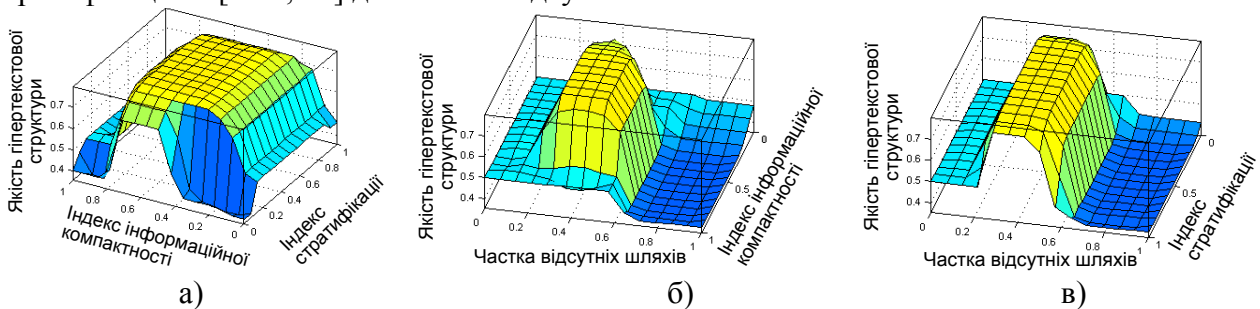


Рисунок 3 – Поверхня “входи-виходи” а) входи – індекс інформаційної компактності, індекс стратифікації; б) входи – частка відсутніх шляхів, індекс стратифікації; в) входи – частка відсутніх шляхів, індекс інформаційної компактності

На рис. 4. наведена загальна схема оптимізації гіпертекстової структури веб-ресурсів, що запропонована у роботі. Першою складовою запропонованої схеми оптимізації є аналіз

гіпертекстової структури, формалізація та оцінювання її якості. Наступна охоплює проведення оптимізації на основі семантичних даних та на основі діяльності користувачів, після чого результати об'єднуються та надаються користувачу системи у вигляді рекомендацій з модифікації гіпертекстових зв'язків Інтернет-ресурсу.

В роботі розроблено метод оптимізації структури сайтів на основі семантичної інформації при такій постановці задачі: для заданої множини ресурсів $S = (s_1, s_2, \dots, s_N)$, сторінок $P = (p_1, p_2, \dots, p_M)$, зв'язків між сторінками $L = (l_1, l_2, \dots, l_K)$ та множини онтологій $O = (o_1, o_2, \dots, o_{|P|})$ здійснити доповнення L згідно з нижченаведеними гіпотезами оптимальної навігації у гіпертексті: гіпертекстовий зв'язок визначає відповідний семантичний зв'язок інформаційних статей, найбільш прийнятним і ефективним способом подання інформації є ієрархічний, в якому вищі рівні ієрархії співвідносяться до нижчих як загальне до окремого.



Рисунок 4 – Загальна схема оптимізації гіпертекстової структури веб-ресурсів

Відповідно до поставленої задачі, запропоновано алгоритм побудови оптимальної структури зв'язків у гіпертексті на основі семантичної інформації виду

1. Визначення множини тематичних кластерів $Y_i = \{g \in G \mid \forall m \in R_i : gIm\}$ для сегмента гіпертекстової мережі на основі методу ідентифікації частих множин ознак, де Y_i – множина ідентифікаторів сторінок, що входять в тематичний кластер i , R – множина ознак кластерів, G – множина об'єктів, M – множина ознак (атрибутів), I – бінарне відношення між G та M ($I \subseteq G \times M$), $g \in G, m \in M$.

2. Побудова формального контексту для кластерів $K_D = (G_D, M_D, I_D)$, де об'єктами є кластери $G_D = (Y_i \mid i=1, \dots, |Y|)$, а ознаками – сукупність усіх ознак об'єктів, що входять у кластер $M_D = \bigcup_{i=1}^{|R|} R_i$, $I_D \subseteq G_D \times M_D$.

3. Встановлення ієрархії відношень між кластерами для контексту K_D шляхом використання властивості часткового порядку на контексті формальних понять для пошуку ієрархічних зв'язків між кластерами.

4. Побудова для сторінок кластера $i = 1, \dots, |Y|$ перетвореної матриці відстаней CDM . Елементами матриці CDM є сторінки $X_{D_i} = (x \mid x \in Y_i)$ з кластера i . На основі перетвореної матриці відстаней та індексу рівновіддаленості вершини визначення центральних вузлів кластерів за метрикою ROC (relative out centrality), індексу центральності веб-сторінки, що показує ступінь досяжності усіх сторінок сегмента гіпертекстової мережі відносно заданої,

$ROC = (ROC_j | j = 1, \dots, |X_{D_i}|)$. Центральною вершиною кластера є сторінка, для якої $ROC_j = \max ROC$.

5. Встановлення ієрархічних зв'язків, визначених на кроці 3 між кластерами, шляхом з'єднання центральних вершин кластерів.

6. Побудова мінімального кістякового дерева між усіма елементами (сторінками) кластера на основі визначеної в роботі характеристики втрат інформаційного перегляду, що залежить від часу перебування користувача на сторінках маршруту та значень центральності переглянутих веб-сторінок згідно з метрикою ROC . При втратах інформаційного перегляду сторінки x_k , рівної p_k , де $x_k \in X_{D_i}$, елементами матриці відстаней будуть p_{km} , а результат повинен задовольняти $\sum_{x_k \in X_{D_i}} p_k \rightarrow \min$.

Відповідно до розробленої в роботі моделі гіпертекстового інформаційного середовища поведінки користувачів є невід'ємною його складовою. Цілі користувачів щодо інформації є різними – для деяких користувачів певний сегмент може бути транзитним на шляху перегляду, для інших – метою пошуку.

Незалежно від цілей, користувачі завжди прагнуть мінімізувати час, необхідний для віднаходження інформації. Проведений аналіз діяльності користувачів у гіпертекстовому середовищі дозволив розробити модель їхньої поведінки, яка ґрунтується на оцінюванні ентропії системи, що дозволило визначити функцію корисності маршруту перегляду інформаційних ресурсів та цілі користувачів.

Для цього в роботі запропонована характеристика втрат інформаційного перегляду (ВІП) інформаційної статті, що враховує основні параметри поведінки користувача в гіпертексті

$$p_i = \frac{\ln(s_i)}{t_i \ln(ROC_i + 2)}, \quad (5)$$

де t_i – середній час перегляду сторінки у припущенні, чим більший час користувач перебуває на сторінці, тим ця сторінка важливіша для досягнення мети пошуку, – якщо $T_{Y_k} = \{T_{ij} | i = Y_k, j = 1..|T_{ij}|\}$ – множина значень часу, витраченого на перегляд сторінок маршруту користувачем k , тоді $t_i = \sum T_{Y_k} / |T_{Y_k}|$, s_i – розмір текстової складової сторінки у байтах.

Модель поведінки користувача в процесі пошуку інформації подана на рис. 5 у вигляді UML-діаграми діяльності, яка відображає характер дій користувача з віднаходження сторінок, які відповідають меті його інформаційного пошуку, витрачаючи більший чи менший час на перегляд деякої сторінки, повертаючись до певних сторінок по декілька разів та перериваючи пошук, якщо він стає надто довгим. Тобто користувач гіпертексту в процесі навігації підсвідомо максимізує корисність інформаційного перегляду

$$u = u(x_1, x_2, x_3, \dots, x_n) \rightarrow \max, \quad (6)$$

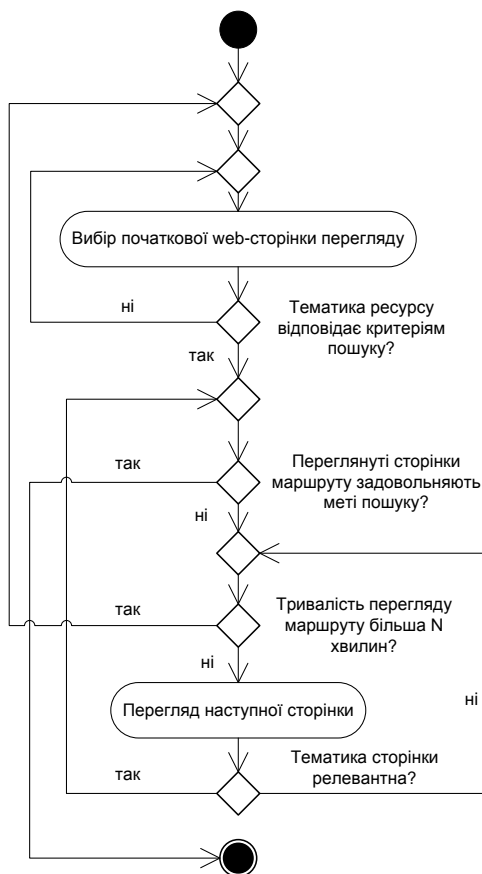


Рисунок 5 – Модель поведінки користувача

де x_1, x_2, \dots, x_n – число переходів на сторінки $1, \dots, n$ у маршруті, за якими користувач здійснює перегляд гіпертексту. Модель системи максимальної корисності з множиною обмежень виду

$$f_k(x_1, x_2, \dots, x_N) \leq g_k, k = 1, \dots, K \quad (7)$$

може бути подана за допомогою функції Лагранжа

$$L = u(x_1, x_2, \dots, x_n) + \sum_k \mu_k (g_k - f_k), \quad (8)$$

де μ_k – множник Лагранжа. Результатом оптимізації, в термінах гіпертекстових систем, є x_i – кількість разів перегляду користувачем сторінки i у заданому гіпертекстовому сегменті.

Для побудови моделі поведінки користувачів застосовано поняття ентропії для гіпертекстових систем як систем із максимальною корисністю, що дозволило звести дану задачу до максимізації ентропії. Для визначення ентропії у випадку гіпертекстової системи було обґрунтовано ВПІ для сторінки у нормованому вигляді $y_i = x_i p_i / I_0$ при обмеженні

$$\sum_i x_i p_i \leq I_0, \quad (9)$$

де I_0 – обмеження на ВПІ перегляду. В загальному вигляді ентропію поведінки користувачів визначено так

$$S = -\sum_i y_i \ln y_i, \quad (10)$$

$$\sum_i y_i \leq 1. \quad (11)$$

В роботі було визначено умову відбору таких сторінок, які сприяють користувачеві в процесі пошуку інформації або відповідають цілям його пошуку

$$p_i < M[P] - \sigma, \quad (12)$$

де p_i – втрати інформаційного перегляду сторінки i , $M[P]$ – математичне сподівання дискретної величини $P = (p_i | i = 1, \dots, n)$, σ – середньоквадратичне відхилення. Тоді V – підмножина P , що містить ВПІ кожної сторінки певного маршруту користувача, кожен елемент якої задовольняє умову (11)

$$V = (v_i | v_i \in P, v_i < M[P] - \sigma, i = 1, \dots, n). \quad (13)$$

На основі (13) визначено поняття збалансованого маршруту, що складається з сторінок, значення ВПІ яких задовольняє обмеження (12). Даний підхід враховує індивідуальні особливості навігації користувачів, що пов'язані з швидкістю перегляду сторінок. Обмеження збалансованого маршруту

$$\sum_{i=1}^N y_i \leq \sum_{j=1}^M y_j. \quad (14)$$

На рис. 6 показана графічна інтерпретація обмеження збалансованого маршруту (14). Приклад, що зображено, відповідає маршруту користувача з 11 переглядів сторінок певного гіпертекстового сегмента. Квадратні позначки відповідають значенням ВПІ кожної сторінки у маршруті. Горизонтальна лінія відповідає значенню математичного очікування P , прямокутник – області, що визначена обмеженням (12) і відповідає сторінкам, що, відповідно до даного обмеження, корелюють із цілями інформаційного пошуку користувачів і формують множину V .

За умов обмежень (9), (11), (14) та ентропії (10), вираз (8) набуває вигляду

$$L = S + \mu_0 \left(1 - \sum_i y_i \right) + \mu_1 \left(\sum_{i=1}^N y_i - \sum_{j=1}^M y_j \right), \quad (15)$$

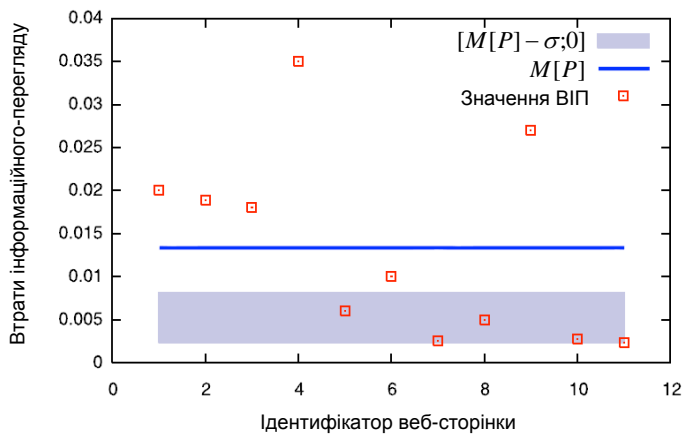


Рисунок 6 – Графічна інтерпретація обмеження збалансованого маршруту

відміну від існуючих, використовує інтелектуальних агентів, які використовують метод оптимізації структури сайтів на основі семантичної інформації, модель гіпертекстового інформаційного середовища і модель поведінки користувачів у гіпертексті.

Запропонована мультиагентна технологія ґрунтується на методі Map-Reduce. Визначені архітектурні особливості системи, типи агентів, особливості їх взаємодії. Встановлена відповідність запропонованих агентів інтелектуальним. Сукупність агентів системи розподіленого обчислення, побудованої на основі мультиагентного підходу, являє собою середовище для виконання обчислювальних задач оптимізації. Основні операції в даній системі виконує інтелектуальний агент, побудований за критеріями інтелектуальності. Агенти розміщуються на веб-ресурсах, на яких необхідно здійснювати оптимізацію, здійснюють автоматичний пошук суміжних ресурсів, що оптимізуються, взаємний обмін інформацією, необхідною для оптимізації. Кількість типів агентів визначається кількістю незалежних операцій, які у комплексі реалізують методуку оптимізації.

Інтелектуальна складова агентів у системі полягає у їх адаптивній поведінці в процесі розв'язання задачі, яка в залежності від завантаження вузлів, може бути частково або повністю делегована іншим агентам даного типу. Відповідно агент, який делегував частину задачі, виконує функції координатора. Архітектура агента підтримує можливість поетапного зниження алгоритмічної складності задачі.

Оскільки мультиагентна система (МАС) передбачає можливість динамічної зміни кількості обчислювальних ресурсів, постає необхідність у визначенні їх кількості для обслуговування потоку запитів. Для цього в роботі запропоновано модель мультиагентної системи, яка базується на теорії масового обслуговування.

У випадку паралельного виконання запиту множиною агентів час, необхідний для виконання запиту, можна визначити як

$$\tau_e = \frac{1}{S_e} \left(\sum_{i=1}^N t(Q_i) + \sum_{i=1}^{N-1} \tau_{i,i+1} \right), \quad (16)$$

де S_e – прискорення процесу виконання певної задачі за рахунок застосування розпаралелювання, $Q_1, Q_2, Q_3, \dots, Q_N$ – продуктивність агентів, N – кількість агентів, які задіяні в обробці запиту користувача.

Показана необхідність використання багатоканальної системи обслуговування типу $G/G/M$ для гарантованого виконання завдання з швидкістю, що не менша за певне заздалегідь встановлене значення, та отримано вираз для розрахунку кількості дублювальних компонентів МАС

$$N_{MAC} = 2\lambda\tau_e, \quad (17)$$

де λ – інтенсивність вхідного потоку запитів МАС. Для управління МАС у її складі передбачений агент, який розподіляє потоки запитів, таким чином забезпечується загальна

та являє собою функцію корисності інформаційного перегляду, яка характеризує поведінку користувача у гіпертексті й поєднує такі характеристики навігації, як час перегляду сторінок, тип сторінок, усі компоненти маршруту.

Запропонована модель враховує індивідуальні характеристики процесу навігації користувача, дозволяє чисельно оцінити ВІП того чи іншого маршруту, а також дає можливість порівняння ВІП різних маршрутів.

У **третьому розділі** запропоновано мультиагентну технологію оптимізації структури Інтернет-ресурсів, яка, на

надійність системи за рахунок дублювання її компонентів та стійкість до перевантажень. Схема взаємодії координаторів виконання задач зображена на рис. 7.

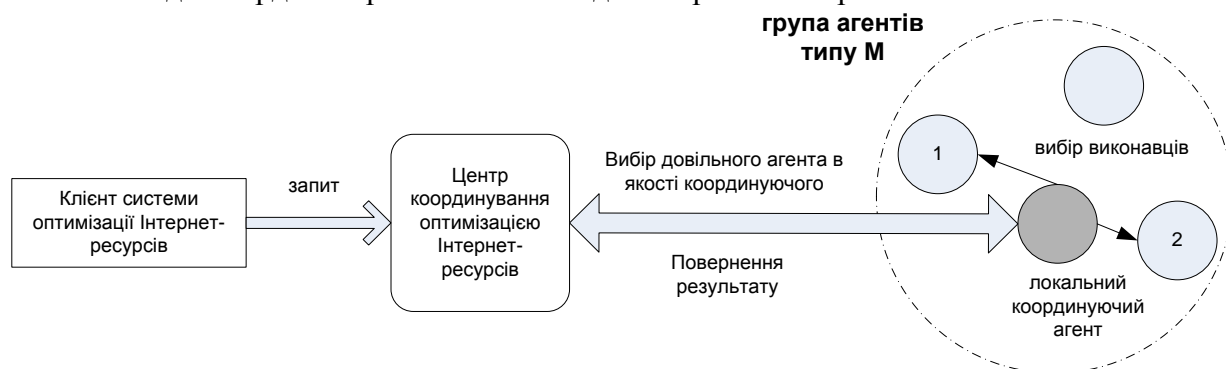


Рисунок 7 – Схема розподілу запитів у мультиагентній системі

Запропонований підхід до декомпозиції задач передбачає використання графу залежності задач (task dependency graph). Для оцінювання прискорення процесу виконання за рахунок розбиття обчислювального алгоритму на незалежні блоки і виконання їх незалежними агентами пропонується використання закону Амдала. У разі неврахування витрат на комунікацію між агентами, прискорення від розподіленого виконання будемо визначати як $S = 1/(\alpha + (1-\alpha)/p)$, де α – частка обчислень, що виконуються послідовно, $1-\alpha$ – частка обчислень, що може бути розпаралелена ідеально, та p – кількість задіяних вузлів.

В роботі пропонується використати пі-числення (pi-calculus) для формалізації процесів взаємодії в розподіленій системі оптимізації Інтернет-ресурсів. Використовуючи відповідні нотації та оператори, в роботі описано основні процеси взаємодії агентів.

Локальна координація агентом обчислювального процесу

$$X = (\nu l)(t).\bar{p}\langle \rangle | p(s_1, s_2, \dots, s_n).\bar{p}\langle t_1, t_2, \dots, t_k, r \rangle | r(r_1, r_2, \dots, r_n).\bar{z}\langle result \rangle, \quad (18)$$

де агент X очікує завдання t по каналу l . При отриманні повідомлення t агент здійснює опитування суміжних агентів шляхом відправки повідомлення у канал p , де паралельний процес агента X очікує на відповіді вільних агентів, що можуть виконати завдання. Після чого агент X здійснює декомпозицію завдання e на частини t_1, t_2, \dots, t_k та відправляє їх агентам s_1, \dots, s_n разом із каналом, на якому агент X очікує на результат. Паралельний процес отримує відповіді по каналу r , об'єднує результат та відправляє центральному координатору по каналу зв'язку z .

Моніторинг зовнішніх об'єктів, стан яких змінюється:

$$Y = (n(x_1, x_2, \dots, x_n)|d(y_1, y_2, \dots, y_n)).[(x_1, x_2, \dots, x_n) \neq (y_1, y_2, \dots, y_n)]\bar{q}(x_1, x_2, \dots, x_n), \quad (19)$$

де n – канал, по якому агент Y отримує дані із зовнішнього середовища; d – канал, по якому агент Y отримує відповідні дані із локальної бази даних. У випадку їх незбігання агент відправляє повідомлення-запит на оновлення тої інформації відповідного ресурсу по каналу q в чергу MAC. Глобальний координатор описується як $Z = q(t).l\langle m \rangle$, що фактично перенаправляє запити із черги q по внутрішньому каналу в систему транспортування повідомлень l . Діяльність звичайного агента, що виконує завдання, описується як $E = (\nu l)(e, x).\varepsilon_l.x\langle m \rangle$, де l – канал, по якому агент отримує завдання t , та канал x , яким необхідно буде відправити

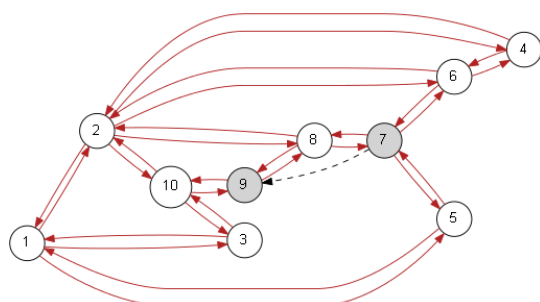


Рисунок 8 – Результат зміни гіпертекстової структури на основі результатів оптимізації функції корисності

результат виконання; ε_i – визначення результату e .

У **четвертому розділі** наведена практична реалізація і застосування інформаційної технології, результати впровадження розробленого програмного забезпечення, використаного для мультиагентної оптимізації Інтернет-ресурсів та здійснена перевірка адекватності та ефективності розроблених в роботі моделей та методів оптимізації.

Перевірку адекватності розробленої моделі максимізації функції корисності, що лежить в основі методу побудови оптимальних маршрутів у гіпертексті, виконано експериментально.

В результаті проведення експерименту були визначені точки входу і точки виходу, між якими виник найбільший потік користувачів та оцінювався час, витрачений користувачами на перегляд їх сторінок.

Відповідно до умови відбору даних (12) було відібрано сторінки, які є найбільш цікавими для користувачів, на основі чого проведено оптимізацію та визначено зв'язки, що їх необхідно змінити.

Для оцінювання адекватності гіпертекстова структура була змінена відповідно до результатів оптимізації та проведено повторне дослідження (рис. 7). Результати аналізу повторного дослідження, проведеного після оптимізації показують зменшення середньої протяжності шляху на 26.2%, підвищення на 20.5% середнього часу перебування користувача на сторінці та як результат, зменшення середнього значення ВПП на 15.4%. Отримані результати підтверджують адекватність розробленої моделі, оскільки в результаті оптимізації було здійснено зниження ВПП та підвищено швидкість досяжності фрагмента гіпертексту.

Розроблене програмне забезпечення дозволяє здійснювати оптимізацію гіпертекстової структури на основі цілей користувачів та семантичної інформації, здійснювати координацію інтелектуальних агентів для задач оптимізації.

Розроблена інформаційна технологія мультиагентної оптимізації Інтернет-ресурсів складається з двох частин – веб-додатка та мультиагентної системи. Веб-додаток дозволяє користувачеві взаємодіяти з МАС, а МАС реалізує всі обчислювальні задачі розроблених методів. Для розробки веб-додатка використано трирівневу архітектуру, що складається з серверної частини, бізнес-логіки та клієнтської складової; для мультиагентної системи використано дворівневу архітектуру, що складається з бізнес-логіки, реалізованої у агентах, систем підтримки їх діяльності та бази даних. Для реалізації МАС була використана бібліотека Jade як засіб для побудови мультиагентних систем загального призначення. Для зберігання онтологій, формалізованих мовою RDF, з метою їх подальшого аналізу використано програмне забезпечення 3store як контейнер метаданих, що підтримує специфікацію SPARQL, та бібліотека Jena для обробки онтологічних даних.

Результати дисертаційних досліджень впроваджені на підприємствах Kamtech (Норвегія, м. Осло), ТОВ “Арісент Україна” (Україна, м.Вінниця), Басейнове управління водних ресурсів річки Південний Буг (Україна, м.Вінниця) та у навчальний процес кафедри КСУ Вінницького національного технічного університету, що підтверджено відповідними актами. Результати впровадження дозволяють зробити висновок про правильність принципових положень, що лягли в основу роботи.

В **додатках** наведено результати моделювання функції корисності перегляду гіпертексту, основні лістинги програмного забезпечення практичної реалізації інформаційної технології, документи та відомості про впровадження результатів, що отримані в дисертації.

ВИСНОВКИ

В результаті виконання дисертаційної роботи розв'язана актуальна наукова задача підвищення швидкості пошуку інформації в мережі Інтернет на основі розробки нової інформаційної технології оптимізації веб-ресурсів. В роботі отримані такі наукові та практичні результати

1. Вперше запропоновано нову модель гіпертекстового середовища, яка поєднує суб'єкти і об'єкти інформаційного пошуку, що дозволило враховувати цілі користувачів при оптимізації гіпертекстової структури веб-ресурсів.

2. Вперше запропоновано метод оптимізації структури сайтів на основі семантичної інформації, який відрізняється функціонуванням в умовах невизначеності відомостей про структуру мережі в цілому, що дозволяє проводити оптимізацію, базуючись лише на даних про окремих та суміжні до нього ресурси.

3. Вперше запропоновано модель поведінки користувачів у гіпертексті, яка ґрунтується на оцінюванні ентропії веб-системи з множиною користувачів, що дозволило визначити функцію корисності маршруту перегляду інформаційних ресурсів та визначити цілі користувачів.

4. Вперше розроблено мультиагентну інформаційну технологію оптимізації гіпертекстових систем, яка, на відміну від існуючих, використовує інтелектуальних агентів, які для здійснення оптимізації використовують метод оптимізації структури сайтів на основі семантичної інформації, модель гіпертекстового інформаційного середовища і модель поведінки користувачів у гіпертексті, що в результаті досліджень дозволило підвищити швидкість пошуку інформації на 15.4%.

5. Розроблено алгоритми ідентифікації цілей користувачів, оптимізації гіпертекстової структури на їх основі, оптимізації гіпертекстової структури на основі семантичної інформації та алгоритм функціонування інтелектуальних агентів і управління їх координацією для оптимізації гіпертекстової структури.

6. Розроблено методичку мультиагентної оптимізації семантичної структури гіпертекстових систем, що використовує розроблені методи, моделі та інформаційну технологію оптимізації Інтернет-ресурсів.

7. Розроблено програмне забезпечення “Multiagent hypertext optimization system” для здійснення процесу оптимізації гіпертекстових систем.

8. Результати дисертаційної роботи впроваджені на підприємствах Kamtech (Норвегія, м. Осло), ТОВ “Арісент Україна” (Україна, м. Вінниця), Басейнове управління водних ресурсів річки Південний Буг (Україна, м. Вінниця) та у навчальний процес кафедри КСУ Вінницького національного технічного університету. Впровадження результатів дисертаційних досліджень підтверджено відповідними актами.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Москвін О. М. Метод захисту інформації в дистанційному курсі / А. С. Васюра, С. М. Москвіна, О. М. Москвін, В. В. Кабачій // Інформаційні технології та комп'ютерна інженерія. – 2007. – №2(13). – С. 199-202.

2. Москвін О. М. Оптимізація структури сайта в умовах неповної інформації [Електронний ресурс] / О. В. Глонь, В. М. Дубовой, О. М. Москвін // Наукові праці Вінницького національного технічного університету. – 2008. – № 1. Режим доступу: http://www.nbu.gov.ua/e-journals/VNTU/2008-1/uk.files/08govoii_uk.pdf.

3. Москвін О. М. Оцінка гіпертекстової структури на основі семантичної інформації / В. М. Дубовой, О. М. Москвін // Вісник Вінницького політехнічного інституту. – 2011. – № 3. – С. 153-157.

4. Москвін О. М. Алгебра узгодження онтологічних контекстів в умовах неповної інформації / В. М. Дубовой, О.М. Москвін // Вісник Вінницького політехнічного інституту. – 2011. – № 5. – С. 244-250.

5. Москвін О. М. Модель поведінки користувача у гіпертексті / В. М. Дубовой, О. М. Москвін // Вісник Вінницького політехнічного інституту. – 2011. – № 6. – С.210-215.

6. Москвін О. М. Інтелектуальна мультиагентна технологія оптимізації Інтернет-ресурсів / В. М. Дубовой, О. М. Москвін // Вісник Вінницького політехнічного інституту. – 2012. – № 5. – С.98-106.

7. Москвін О. М. Розробка нечіткої системи класифікації гіпертекстових структур [Електронний ресурс] / В. М. Дубовой, О. М. Москвін // Наукові праці ВНТУ. – 2009. – № 2. – Режим доступу до журн.: http://www.nbuu.gov.ua/e-journals/vntu/2009-2/2009-2.files/uk/09vmdfcs_ua.pdf
8. Москвін О. М. Синтез інформаційної моделі CD-CASE технології для систем дистанційного навчання / С. М. Москвіна, О. М. Москвін, В. В. Кабачій, Л. П. Погонєць // Інтернет-Освіта-Наука, П'ята міжнародна конференція ІОН-2006, Вінниця. – 2006, 10-14 жовтня. Збірник матеріалів конференції. Т. 1. Вінниця: УНІВЕРСУМ-Вінниця, 2006. – С. 155-158.
9. Москвін О. М. Оптимізація структури сайту в умовах неповної інформації на основі використання мурашиних алгоритмів [Електронний ресурс]: Дев'ята міжнародна наукова конференція “Контроль і управління в складних системах”, Вінниця. – 2008, 21-24 жовтня. / В. М. Дубовой, О. М. Москвін // Тези доповідей. – 2008. – Режим доступу до журн.: http://mccs.vntu.edu.ua/mccs2008/materials/subsection_2.2.pdf
10. Москвін О. М. Оцінка гіпертекстової структури на основі семантичної інформації / В. М. Дубовой, О. М. Москвін // 17 Міжнародна наукова конференція з автоматичного управління “Автоматика-2010”, Харків. – 2010, 27-29 вересня. Тези доповідей. Харків: – ХНУРЕ. – Т. 2. – С. 101-102.
11. Москвін О. М. Аналіз семантичної інформації гіпертексту / В. М. Дубовой, О. М. Москвін // Міжнародна науково-практична конференція “Інформаційні технології та комп'ютерна інженерія”. Вінниця. – 2010, 19-21 травня. Тези доповідей. – Вінниця, 2010. – С. 124-125.
12. Москвін О.М. Метод оптимізації гіпертекстових ресурсів на основі семантичної інформації / В.М. Дубовой, О.М. Москвін // Десята міжнародна наукова конференція “Контроль і управління в складних системах”. Вінниця. – 2010, 19-21 жовтня. Тези доповідей. Вінниця. – 2010. – С.39.
13. Москвін О.М. Модель поведінки користувача у гіпертекстовому середовищі / О.М. Москвін, В.М. Дубовой // Міжнародна наукова конференція “Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту”. – Євпаторія. – 2012, 27-31 травня. Збірка наукових праць. Херсон: ХНТУ, 2012. – С. 133.
14. Москвін О. М. Інтелектуальна мультиагентна технологія оптимізації Інтернет-ресурсів / В. М. Дубовой, О. М. Москвін // Третя міжнародна науково-практична конференція “Інформаційні технології та комп'ютерна інженерія”. Вінниця. – 2012, 29-31 травня. Тези доповідей. – Вінниця, 2012. – С. 22-23.
15. Москвін О. М. Комп'ютерна програма “Автоматизована мультиагентна система керування розподіленим обчислювальним процесом оптимізації структури Інтернет-ресурсів” / В. М. Дубовой, О. М. Москвін // Свідоцтво про реєстрацію авторського права на твір № 42223. – К.: Державний департамент інтелектуальної власності України. — Дата реєстрації: 13.02.2012 р.
16. Москвін О. М. Комп'ютерна програма “Автоматизована система оптимізації структури Інтернет-ресурсів” / В. М. Дубовой, О. М. Москвін // Свідоцтво про реєстрацію авторського права на твір № 42225. – К.: Державний департамент інтелектуальної власності України. – Дата реєстрації: 13.02.2012 р.

АНОТАЦІЯ

Москвін О. М. Інформаційна технологія мультиагентної оптимізації Інтернет-ресурсів. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Вінницький національний технічний університет, Вінниця, 2012.

Метою роботи є підвищення швидкості пошуку інформації в мережі Інтернет на основі розробки нової інформаційної технології мультиагентної оптимізації веб-ресурсів. Запропоновано метод оптимізації структури сайтів на основі семантичної інформації, який відрізняється функціонуванням в умовах невизначеності відомостей про структуру мережі в цілому. Запропоновано модель гіпертекстового інформаційного середовища, яка поєднує суб'єкти і об'єкти інформаційного пошуку, що дозволило враховувати цілі користувачів при оптимізації досяжності інформації. Вперше запропоновано модель поведінки користувачів у гіпертексті, яка ґрунтується на оцінюванні ентропії веб-системи з множиною користувачів.

На основі запропонованих моделей, методів і алгоритмів розроблено інформаційну технологію мультиагентної оптимізації гіпертекстових систем, яка, на відміну від існуючих, використовує інтелектуальних агентів. Створено програмне забезпечення "Multiagent hypertext optimization system" для здійснення процесу оптимізації гіпертекстових систем. Ключові слова: гіпертекст, гіпертекстова система, Інтернет-ресурс, модель поведінки користувача, інформаційний пошук, інформаційна цінність, мультиагентна технологія, розподілені обчислення, оптимізація структури сайтів, семантична інформація.

АННОТАЦІЯ

Москвин А. М. Информационная технология мультиагентной оптимизации Интернет-ресурсов. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических по специальности 05.13.06 – информационные технологии. – Винницкий национальный технический университет, Винница, 2012.

Диссертация посвящена разработке новой информационной мультиагентной оптимизации Интернет-ресурсов с целью повышения скорости достижимости информации в гипертекстовых системах.

Проведен анализ существующих подходов к решению проблемы низкой скорости достижимости информации в гипертекстовых системах, на основе которого выбрано направление и определены задачи исследования. Предложена модель гипертекстовой информационной среды, которая объединяет субъектов и объектов информационного поиска. Разработанная модель дает возможность перейти от анализа зависимостей гипертекстовых ссылок к анализу предметной области информации представленной гипертекстом, как с точки зрения семантики, так и информации, которая является целью поиска пользователей. Также модель представляет собой основу для проведения анализа и оптимизации ссылочной структуры Интернет-ресурсов на основе поведения пользователей, которое рассматривается с точки зрения информационных доменов, описанных средствами Semantic Web.

Предложен метод оптимизации структуры сайтов на основе семантической информации, который отличается функционированием в условиях неопределенности сведений про общую структуру сети, что позволяет проводить оптимизацию, основываясь лишь на данных про структуру отдельного ресурса и смежных ресурсов.

Предложена модель поведения пользователей в гипертексте, основанная на оценивании энтропии веб-системы с множеством пользователей, что позволило определить функцию полезности маршрута просмотра информационных ресурсов и целей пользователей. Модель объединяет такие характеристики навигации, как время просмотра страниц, все компоненты маршрута, энтропию. Учет вышеуказанных параметров позволяет также учитывать индивидуальные характеристики процесса навигации пользователя.

Результаты экспериментальных исследований применения модели поведения пользователей для оптимизации гипертекстовой структуры позволили увеличить скорость поиска информации на 15.4% в тестовой сети ресурсов.

На основе предложенных моделей, методов и алгоритмов разработана информационная технология мультиагентной оптимизации гипертекстовых систем, которая, в отличие от существующих, использует интеллектуальных агентов.

Разработано программное обеспечение “Multiagent hypertext optimization system” для совершения процесса оптимизации гипертекстовых систем, которое позволяет оптимизировать гипертекстовую структуру на основе целей пользователей и семантической информации, производить управление координацией интеллектуальных агентов для задач оптимизации.

Результаты диссертационной работы внедрены на предприятиях Kamtech (Норвегия, г. Осло), ООО “Арисент Украина”, Бассейновое управление водных ресурсов реки Южный Буг (Украина, г. Винница) и в учебный процесс кафедры компьютерных систем управления Винницкого национального технического университета, что подтверждено соответствующими актами. Результаты внедрения позволяют сделать вывод о правильности принципиальных рассуждений, которые легли в основу работы.

Ключевые слова: гипертекст, гипертекстовая система, Интернет-ресурс, модель поведения пользователя, информационный поиск, информационная ценность, мультиагентная технология, распределенные вычисления, оптимизация структуры сайтов, семантическая информация.

ANNOTATION

Moskvin O.M. Information technology of multiagent Internet-resources optimization. – Manuscript.

Thesis for a candidate's degree by speciality 05.13.06 – information technologies. – Vinnytsia National Technical University, Vinnytsia, 2013.

Goal of the work lies in increasing the speed of information location in Internet by developing new information technology of multiagent web-resource optimization. Proposed method of web site hypertext structure optimization bases on semantic information, which differs by operating in conditions of uncertainty about whole network structure, which allows fulfilling optimization process being aware of only destination web site structure and adjacent to it ones. Proposed model of hypertext information environment that joins subjects and objects of the information search process, that allowed to determine user goals during attainability optimization. Proposed user behavior model in hypertext that's based on system entropy evaluation that allowed to determine view utility function and user goals for information retrieval process.

Based on proposed models, methods and algorithms, developed information technology for multiagent hypertext systems optimization, that differs in using intellectual agents. “Multiagent hypertext optimization system” software application has been developed for fulfilling hypertext systems optimization process.

Key words: hypertext, hypertext system, Internet-resource, user behavior model, information search, information value, multiagent technology, distributed computing, web site structure optimization, semantic data.

Підписано до друку 15.03.2013 р. Формат 29,7×42 $\frac{1}{4}$

Наклад 120 прим. Зам. №

Віддруковано в комп'ютерному інформаційно-видавничому центрі
Вінницького національного технічного університету
Вінниця, вул. Хмельницьке шосе, 95. Тел.: 59-81-59