

НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ
ИНСТИТУТ ПРОБЛЕМ РЕГИСТРАЦИИ ИНФОРМАЦИИ НАН УКРАИНЫ

НАЦИОНАЛЬНАЯ АКАДЕМИЯ ПРАВОВЫХ НАУК УКРАИНЫ
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ ИНФОРМАТИКИ И
ПРАВА

НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ УКРАИНЫ
«КИЕВСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ»

УЧЕБНО-НАУЧНЫЙ ЦЕНТР ИНФОРМАЦИОННОГО ПРАВА И
ПРАВОВЫХ ВОПРОСОВ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ
ФАКУЛЬТЕТА СОЦИОЛОГИИ И ПРАВА

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И БЕЗОПАСНОСТЬ

МАТЕРИАЛЫ XV МЕЖДУНАРОДНОЙ
НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ

ВЫПУСК 15

Киев – 2015

*Рекомендовано к печати ученым советом
Института проблем регистрации информации НАН Украины
(протокол № 12 от 15 сентября 2015 г.)*

**Информационные технологии и безопасность. Материалы XV
Международной научно-практической конференции ИТБ-2015.** – К.:
ИПРИ НАН Украины, 2015. – 250 с. ISBN: 978-966-2344-45-5

В сборник вошли материалы докладов, представленных на XV
Международной научно-практической конференции «Информационные
технологии и безопасность» (ИТБ-2015, 21 октября 2015 года, г. Киев,
Украина).

В сборнике представлены статьи, посвященные вопросам внедрения
информационных технологий, актуальным проблемам обеспечения
информационной и кибербезопасности, противодействия информационной
агрессии и кибертерроризму, проведения информационно-аналитических
исследований на основе контента сети Интернет, правового обеспечения
информационной безопасности.

Для специалистов в области информационных технологий,
информационной безопасности, информационного права, а также для
аспирантов и студентов старших курсов высшей школы соответствующих
специальностей.

Редакционная коллегия:

*А.Г. Додонов, д.т.н., профессор; В.Г. Пилишчук, д.ю.н., профессор, член-корр.
НАПрН Украины; А.М. Богданов, д.т.н., профессор; Д.В. Ландэ, д.т.н.,
с.н.с.; В.В. Мохор, д.т.н., профессор; Н.А. Ожеван, д.ф.н., профессор; В.Н.
Фурашев, к.т.н., с.н.с.; Е.С. Горбачик, к.т.н., с.н.с.; М.Г. Кузнецова, к.т.н.,
с.н.с.*

ISBN 978-966-2344-45-5

- © Институт проблем регистрации информации НАН Украины, 2015
- © Научно-исследовательский институт информатики и права НАПрН Украины, 2015
- © Учебно-научный центр информационного права и правовых вопросов информационных технологий ФСП НТУУ «КПИ», 2015
- © Коллектив авторов

АЛГОРИТМИ ПОБУДОВИ ОНТОЛОГІЇ НА ОСНОВІ ЗОНДУВАННЯ НАУКОВОЇ ЛІТЕРАТУРИ

*Ланде Д.В., Прищета С.В., Синькова Т.В.
Інститут проблем рестрації інформації НАН України*

Побудова великої галузевої онтології, як правило, складна науково-практична проблема [1], [2]. Основа цього процесу – побудова термінологічної основи і визначення деяких семантичних зв'язків [3].

Зондування наукометричної мережі

У рамках даної роботи розглядається метод побудови моделі предметної області (ПрО), заснований на знаннях, закладених у спеціалізованих текстах, що формуються фахівцями, експертами, авторами наукових праць.

Будемо розглядати зондуванням великих інформаційних мереж – таких, що з технологічних причин не підлягають повному скануванню, як вибірку невеликого обсягу найважливішого змісту. Для цього оберемо мережу тегів наукометричного сервісу Google Scholar Citations, вузли якої – поняття, марковані тегами, а ребра – деякі семантичні зв'язки між ними, обумовлені суміжними інтересами окремих авторів.

Як теги позначимо поняття наукового напрямку. Після введення відповідного тегу до адресного рядку, в інтерфейсі сервісу Google Scholar Citations у ранжируваному вигляді посторінково відображаються імена вчених, які позначили свої наукові інтереси заданим поняттям. Для кожного вченого вказані також інші поняття, що входять до сфери його наукових інтересів. Наприклад, для автора Piek Vossen визначені ще такі теги, як Lexical Semantics, Word-Sense-Disambiguation, Event Extraction, Text Mining, Knowledge Acquisition. Множина тегів-понять утворює мережу, похідну від біграфа «вчений-поняття» (біграф – це графа, множини вершин якого можна розбити на дві частини таким чином, що кожне його ребро з'єднує якусь вершину з однієї частини з якоюсь вершиною іншої частини). В рамках даної роботи як одну частину можна розглядати множини вчених, а іншу – множини тегів). Для моделі ПрО будемо розглядати похідну від цієї мережі, а саме мережу зв'язків окремих понять через авторів, яким ці поняття приспані. Очевидно, зв'язки такої «похідної» мережі можуть мати вагу, пропорційну кількості авторів, яким приписується відповідна пара понять.

Іноколи, коли теги зазначені окремими вченими можуть належати до різних галузей науки, однак, попередньо проведені дослідження показують, що на невеликий, але досить репрезентативній вибірці (порядку сотні тегів), невелика частота нетематичних тегів забезпечує їх автоматичне «відсіювання» за деяким невеликим порогом.

Модель зондування мережі

Зондування модельної мережі здійснюється за принципом, подібним до того, що застосовується при пошуку інформаційних ресурсів в пінгвіних мережах (від англ. Peer-to-peer, мережа – однорангова мережа, заснована на рівноправності учасників) [4], [5], [6].

Алгоритм зондування мережі Google Scholar Citations

Алгоритм, який застосовувався до модельних мереж, був адаптований до реальної мережі тегів сервісу Google Scholar Citations наступним чином:

1. Експертним шляхом визначається невеликий перелік базових тегів (ключових слів, відповідних найбільш важливим поняттями ПрО).
2. Вибирається тег із визначеного експертами переліку.
3. Відкриваються сторінки веб-сервісу, що відповідають цьому тегу (максимальна кількість таких сторінок параметрично обмежується задалегідь).
4. До створеної мережі додаються всі теги, що містяться на вибраних сторінках (сусідні теги).
5. Із сусідніх тегів вибирається той, на відповідні сторінки сервісу якого планується перейти для подальшого аналізу. Цей тег з найбільшим ступенем серед сусідніх тегів, який також задовольняє тематиці обраної ПрО і не входить до складу тих тегів, до сторінок яких вже було здійснено перехід.
6. Якщо такий тег обрано, то відбувається перехід до пункту 3.
7. Якщо такого тегу не існує, але перелік базових тегів не завершено, то здійснюється перехід до наступного базового тегу з початкового переліку, тобто перехід до пункту 2. У іншому випадку вважається, що мережа зондування побудована.

Процес зондування мережі, починаючи з певного вузла, припиняється при «зациклованні», тобто коли відповідно до алгоритму відбувається перехід до вже пройденого тегу, а також при відхиленні від основної тематики сусідніх тегів, що залишилися.

Відхилення від основної тематики визначається експертами при автоматизованому зондуванні або з урахуванням лексичного складу тегів при повністю автоматичному скануванні. У разі автоматичного виконання алгоритму виконується обмеження за допомогою так званих «плюс» і «стоп-словників» – наборів спеціальних шаблонів («плюс» і «стоп-словник» в рамках даної роботи – набори шаблонів, підрядків, які повинні обов'язково входити або, відповідно, не входити в рядки, відповідні тегам). При цьому саме «зацикловання» є ознакою переходу до наступного базового тегу або завершення процесу зондування.

Експертами з ПрО виконується формування стартового переліку вузлів-понять і правил відбору «кінцевих» вузлів. Для побудови моделі ПрО (в розглянутому прикладі для області екстрагування подій із текстів) експертним шляхом були визначені базові теги англійською мовою: Event

знання, закладені самими авторами публікацій, теги відмічені ними як головні. Тобто експертне середовище в цьому випадку істотно розширюється.

Модель застосована для напрямку досліджень «екстрагування понять», але її можна використовувати і для інших наукових областей. Зокрема, вже побудовані подібні мережі для напрямків штучного інтелекту, багатогентних системи і складних мереж (Complex Networks).

1. *Добров Б.В.* Онтологии и тезаурусы / Добров Б.В., Соловьев В.Д., Лукашевич Н.В., Иванов В.В. // Модели, инструменты, приложения. Бином, 2009. – 173 с.
2. *Ландо Д.В.* Подход к созданию терминологических онтологий / Ландо Д.В., Снарский А.А. // Онтология проектирования, 2014. – № 2(12). – С. 83-91.
3. *Чанышев О.Г.* Автоматическое построение терминологической базы знаний / Чанышев О.Г. // Труды 10-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL'2008, Дубна, Россия, 2008. – С. 85-92.
4. *Zeinalipour-Yazti D.* Information Retrieval in Peer-to-Peer Networks / Zeinalipour-Yazti D., Kalogeraki V., Gunopulos D // IEEE CISE Magazine, Special Issue on Web Engineering, 2004. – pp. 1-13.
5. *Kalogeraki V.* A Local Search Mechanism for Peer-to-Peer Networks / Kalogeraki V., Gunopulos D., Zeinalipour-Yazti D. // Proc. of CIKM'02, McLean VA, USA, 2002.
6. *Yang B.* Efficient Search in Peer-to-Peer Networks / Yang B., Garcia-Molina H. // Proc. of ICDCS'02, Vienna, Austria, 2002.
7. *Erdős, P.* On The Evolution of Random Graphs / Erdős, P., Rényi, A. // Magyar Tud. Akad. Mat. Kutató Int. Közl. 5, 1960. – 17-61 pp.
8. *Albert* Statistical mechanics of complex networks / Albert, Réka; Barabási, Albert-László // Reviews of Modern Physics 74, 2002. – 47–97 pp.
9. *Ландо Д.В.* Моделирование контентных сетей // Проблемы информатизації та управління: Збірник наукових праць: Випуск 1

Зубок В.Ю.

ОПЦІОНОВАННЯ СТАНУ МЕРЕЖІ ІНТЕРНЕТ СТОСОВНО ЗАГРОЗ ЗМІНИ ЇЇ ТОПОЛОГІЇ..... 100

Ірха Ю.Б.

ЕКСТРЕМІСТСЬКА ДІЯЛЬНІСТЬ У МЕРЕЖІ ІНТЕРНЕТ: ПРАВОВІ ПРОБЛЕМИ ВИЯВЛЕННЯ ТА ПРОТІДІЇ В УКРАЇНІ..... 106

Корж І.

ПРОБЛЕМНІ АСПЕКТИ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ЛЮДИНИ, СУСПІЛЬСТВА І ДЕРЖАВИ..... 110

Косоогов О.М.

ДЕРЖАВНА СИСТЕМА ІНФОРМАЦІЙНОГО ПРОТИБОРСТВА В ОСОБЛИВИХ УМОВАХ..... 115

Красноступ Г.

ПРОЗОРІСТЬ ВЛАСНОСТІ У СФЕРІ ТЕЛЕРАДІОМОВЛЕННЯ: ПРИОРТЕТНИЙ НАПРЯМ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ..... 117

Крючин А.А., Беляк Е.В., Морозов Е.М.

ОПТИЧЕСКИЕ НОСИТЕЛИ ИНФОРМАЦИИ С ПОЛЯРИЗАЦИОННЫМ ШИФРОВАНИЕМ ДАННЫХ..... 122

Кузнецова Н.В.

ПИТАННЯ ЯКОСТІ, КОНФІДЕНЦІЙНОСТІ ТА КОРЕКТНОСТІ ДАНИХ В ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЯХ АНАЛІЗУ ФІНАНСОВИХ РИЗИКІВ..... 131

Кузнецова М.Г.

МЕХАНІЗМИ ПІДВИЩЕННЯ ЖИВУЧОСТІ В СИСТЕМАХ ОРГАНІЗАЦІЙНОГО УПРАВЛІННЯ..... 137

Кузьмичев А. П.

ОПТИМИЗАЦИОННОЕ МОДЕЛИРОВАНИЕ В EXCEL ЗАДАЧ О ПОТОКАХ В КОММУНИКАЦИЯХ..... 141

Д.В. Ланде, С.В. Прищепя, Т.В. Силькова

АЛГОРИТМИ ПОБУДОВИ ОНТОЛОГІЇ НА ОСНОВІ ЗОНДУВАННЯ НАУКОВОЇ ЛІТЕРАТУРИ..... 149

Луж'яничук Р.В.

ДЕЯКІ ПИТАННЯ ДЕРЖАВНОГО РЕГУЛЮВАННЯ У СФЕРІ ЗАБЕЗПЕЧЕННЯ КІБЕРБЕЗПЕКИ: ЗАРУБІЖНИЙ ДОСВІД..... 154

Маріц Д.

"КІБЕРАТАКА" ЯК СПОСІБ ВЧИНЕННЯ ЗЛОЧИНІВ В ІНФОРМАЦІЙНОМУ ПРОСТОРІ..... 158