

# Прилагане на количествени оценки при информационно търсене на документи

---

**БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ  
ИНСТИТУТ ПО ИНФОРМАЦИОННИ И КОМУНИКАЦИОННИ  
ТЕХНОЛОГИИ**

Светломир Минков Станчев

ДИСЕРТАЦИЯ

за придобиване на образователната и научна степен „доктор“  
по специалност 02.21.10 Приложение на принципите и методите на  
кибернетиката в различни области на науката

Научен ръководител: проф. д-н Тодор Стоилов

- 
- Начало на докторантурата 01.01.2009
  - Край на докторантурата 31.12.2013
  
  - Взети изпити:
    - Базов специализиран предмет
    - Интернет технологии за управление
    - Проектиране на интернет приложения
    - Компютърни умения - Java
    - Английски език
-

---

Тема на дисертационния труд:

Прилагане на количествени оценки  
при информационно търсене на  
документи

---

# Публикации

---

- Семантични мрежи и тяхното приложение в изграждане на уеб приложения, сп. Автоматика и информатика брой 1 Съюз по автоматика и информатика, 2012 28-33 стр.
  - Модели информационно търсене, Доклад МЕЖДУНАРОДНА КОНФЕРЕНЦИЯ АВТОМАТИКА И ИНФОРМАТИКА'12 3 – 5 октомври 2012 г. София, България Съюз по автоматика и информатика.
-

# Участие в проекти

---

- **№ BG051PO001-3.3.06-0048** ИЗГРАЖДАНЕ И РАЗВИТИЕ НА МЛАДИ ВИСОКОКВАЛИФИЦИРАНИ ИЗСЛЕДОВАТЕЛИ ЗА ЕФЕКТИВНО ПРИЛАГАНЕ НА БИОМЕДИЦИНСКИТЕ ИЗСЛЕДВАНИЯ ЗА ПОДОБРЯВАНЕ КАЧЕСТВОТО НА ЖИВОТ
-

# Цели

---

Ще бъде разработен алгоритъм за информационно търсене с използването на семантични езикови модели.

Ще бъде използвана количествена оценка на семантичното разстояние, за да бъдат проектирани и експериментирани алгоритми за информационно търсене в неструктурирани масиви от символна информация.

Целта е създаване на семантичен езиков модел, който е аналог на статистическия езиков модел, като разликата между тях е промененото вероятно разпределение в модела с добавена семантична оценка на вероятността за срещаните понятия.

Чрез този семантичен езиков модел ще бъде измерено реалното вероятно разпределение на понятията в този семантичен езиков модел, което ще позволи да се намерят документи със семантично съдържание, по близко до това на запитването.

---

# Предлаган алгоритъм за информационно търсене

---

1. Информационно търсене чрез традиционна търсеща машина
  2. Подготвителни стъпки за построяване на семантичен езиков модел на всеки документ, резултат от традиционното търсене.
  3. Построяване на семантичен езиков модел на всеки документ от списъка с резултати
  4. За всеки документ намиране на вероятността документа и заявката да имат един и същ семантичен езиков модел.
  5. Класиране на документите в съответствие с техните вероятности от стъпка 4
-

# Дисертационен труд

---

- Глава 1 Анализ на състоянието на проблема завършена
  - Глава 2 Количествени критерии за оценка на семантичната близост на документи 50% завършена
  - Глава 3 Проектиране на семантичен езиков модел. 25% завършена
  - Глава 4 Експериментални изследвания
-



# Планирани етапи

---

1. Допълнителни проучвания в областта на статистическите езикови модели и алгоритмите за намиране на семантичната зависимост между понятия.
  2. Имплементиране на описания алгоритъм за информационно търсене чрез семантични езикови модели.
  3. Експериментални резултати
-



Благодаря за вниманието

