

# Chavdar V. Botev

Ph.D. Student  
Computer Science Department  
Cornell University

4130 Upson Hall  
Ithaca, NY 14853  
(+1 607) 262 0694, 255 9211  
cbotev@cs.cornell.edu  
<http://www.cs.cornell.edu/~cbotev/>

## Thesis

|                     |   |
|---------------------|---|
| Title               | Integration of Structured and Full-Text Search in XML |
| Expected graduation | May 2008  |

## Research Interests

- XML databases
- Information retrieval
- Information integration

## Education

|                |  |
|----------------|--|
| 2002 – present | Ph.D. student in Computer Science at the Computer Science Department, Cornell University |
| 1995 – 2000    | B.Sc./M.Sc. in Informatics at Sofia University, Sofia, Bulgaria (GPA 5.9 out of 6)       |

## Professional Experience

|   |   |
|---|---|
| Spring 2007, Spring 2006                              | TA for CS 530 (The Architecture of Large-Scale Information Systems), Cornell University   |
| Sep 2006 – Dec 2006                                   | Intern at Yahoo! Research<br>I worked on a platform for building data-driven web applications by non-programmers.   |
| May 2004 – Aug 2004                                   | Summer intern at IBM Almaden Research Center, Trevi research group<br>I worked on different inverted-list organizations for Trevi, the IBM intranet search engine. I studied static rank bucketing in the inverted lists and its effects on query results.  |
| Summer 2002   | RA of Prof. J. Shanmugasundaram, Cornell University   |
| Spring 2003 – Fall 2005                               | I worked on several projects. I helped with the development of XRank – a link-analysis-based rank for XML documents collections.  |
| Summer 2006,<br>Summer 2007,<br>Fall 2007-Spring 2008 | I am involved in the Quark project, developing a semi-structured database engine. I work on extending it with full-text search capabilities. This work is in conjunction with my work on TeXQuery, a full-text search extension for XQuery. I have also studied query and scoring algorithms for context-sensitive search in XML databases. |
| Fall 2002   | TA for CS 432 (Database Management Systems), Cornell University   |
| 1999 – 2000   | Cornell University TA Award<br>TA for Artificial Intelligence, Sofia University   |

|             |   |
|-------------|---|
| 1998 – 2002 | Managing Director, BMI Ltd, Bulgaria<br>I developed commercial software for a retailer. |
| 1997 – 1999 | TA for Data Structures, Sofia University  |
| 1996 – 1998 | Programmer, VBMI ET, Bulgaria<br>I developed commercial software for a retailer.        |

## Professional Services

- Invited expert at the Full-Text Task Force of the W3C XQuery Working Group, editor of the XQuery 1.0 and XPath 2.0 Full-Text proposal
- External Reviewed, VLDB 2007
- External Reviewer, ICDE 2006
- External Reviewer, ASIAN 2005
- External Reviewer, Comad 2005
- External Reviewer, ICDE 2005
- External Reviewer, SIGMOD 2005
- External Reviewer, WebDB 2005
- External Reviewer, XSym 2005

## Publications

- F. Shao, L. Guo, C. Botev, A. Bhaskar, M. Chettiar, F. Yang, J. Shanmugasundaram, "Efficient Keyword Search over Virtual XML Views", VLDB Conference, 2007.
- S. Amer-Yahia, C. Botev, J. Doerre, J. Shanmugasundaram, "XQuery Full-Text Extensions Explained", IBM Systems Journal 45(2), 2006.
- C. Botev, S. Amer-Yahia, J. Shanmugasundaram, "Expressiveness and Performance of Full-Text Search Languages", EDBT Conference, March 2006.
- C. Botev, J. Shanmugasundaram, "Context-Sensitive Keyword Search and Ranking for XML", Workshop on the Web and Data Bases (WebDB), June 2005.
- S. Amer-Yahia, C. Botev, S. Buxton, P. Case, J. Doerre, D. McBeath, M. Rys, J. Shanmugasundaram. "XQuery 1.0 and XPath 2.0 Full-Text". W3C Working Draft, 09 July 2004.
- S. Amer-Yahia, C. Botev, J. Robie and J. Shanmugasundaram. "TeXQuery: A Full-Text Search Extension to XQuery". WWW Conference, 2004.
- L. Guo, F. Shao, C. Botev, J. Shanmugasundaram. "XRANK: Ranked Keyword Search over XML Documents". SIGMOD, San Diego, 2003.
- C. Botev. "A Design of a Web-Based Cataloguing Information System". Master's Thesis (in Bulgarian), Sofia University, 2000.

## Demos

- A. Bhaskar, C. Botev, M. Muthaia Chettiar, L. Guo., J. Shanmugasundaram, F. Shao, F. Yang. "Quark: An Efficient XQuery Full-Text Implementation", SIGMOD Conference, June 2006.
- Chavdar Botev, Nadav Eiron, Marcus Fontoura, Ning Li, Eugene J. Shekita. "Static score bucketing in inverted indexes", CIKM Conference, 2005.
- C. Botev, S. Amer-Yahia, J. Shanmugasundaram. "A TeXQuery-Based XML Full-Text Search Engine", SIGMOD Conference, 2004.

## **Development Experience**

- Primary programming languages: Java, J2EE, C/C++, Delphi
- Web development: HTML, Dynamic HTML, JavaScript, AJAX, XSLT
- Data query languages: SQL, XQuery/XPath, XQuery Full-Text
- Platforms: Windows, Linux