Vasil Slavov

+1-816-255-0338 vgslavov@umkc.edu http://about.me/vgs

Research Interests

- Linked Data, Semantic Web, Big Data, NoSQL systems
- Scalable RDF query processing using a cloud infrastructure, RDF quadruples
- Federated SPARQL query processing
- Cardinality estimation of queries over semi-structured data using gossip-based algorithms
- RDF and XML data models in DHTs

Education

University of Missouri-Kansas City, Kansas City, MO
Ph.D. Candidate in Computer Science, GPA 3.95 **Spring 2011 – December 2014**

- Dissertation topic: Fast Processing of SPARQL Queries on RDF Quadruples
- Advisor: Professor Praveen Rao

M.S. in Computer Science, GPA 3.95 **Spring 2008 – Summer 2012**

- Thesis topic: A Study of Gossip Algorithms for Internet-Scale Cardinality Estimation of Distributed XML Data
- Advisor: Professor Praveen Rao

William Jewell College, Liberty, MO
B.A. in Computer Science & Math, Summa Cum Laude, GPA 3.97 **Spring 2001 – Spring 2005**

Honors and Awards

- Outstanding Ph.D. Student in CS, 2012, School of Computing and Engineering, UMKC
- Balaji Krithikaivasan Memorial Student Travel Grant, CSEE, 2012, 2014, UMKC
- Dean's International Computing & Engineering (DICE) award, SCE, UMKC
- Computer Science Faculty Award for the Outstanding Junior and Senior in CS, 2004, 2005, WJC
- Phi Epsilon Honor Society, 2005
- Who's Who Among Students in American Universities and Colleges, 2005
- Kappa Mu Epsilon National Mathematics Honor Society, 2003
- Alpha Lambda Delta National Academic Honor Society for Freshmen, 2001
- Dean's List (7 semesters), WJC

Publications

Vasil Slavov, Anas Katib, Praveen Rao, Srivenu Paturi, Dinesh Barenkala. "Fast Processing of SPARQL Queries on RDF Quadruples." To appear in *Proceedings of the 17th International Workshop on the Web and Databases* (**WebDB 2014**), Snowbird, UT, June 2014. `http://v.web.umkc.edu/vsfgd/files/riq-webdb-paper.pdf`

Vasil Slavov, Anas Katib, and Praveen Rao. "Tool for Internet-Scale Cardinality Estimation of XPath Queries over Distributed Semistructured Data." *Proceedings of the 30th IEEE International Conference on Data Engineering* (**ICDE 2014**), Chicago, IL, April 2014. (Demo) `http://dx.doi.org/10.1109/ICDE.2014.6816758`

Vasil Slavov and Praveen Rao. "A Gossip-Based Approach for Internet-Scale Cardinality Estimation of XPath Queries over Distributed Semistructured Data." *The International Journal on Very Large Databases* (**VLDB Journal 2014**), Volume 23, Issue 1, 2014, pp. 56–71. `http://dx.doi.org/10.1007/s00778-013-0314-1`

Vasil Slavov, Praveen Rao, Srivenu Paturi, Tivakar Swami, Michael Barnes, Deepthi Rao, and Raghuvarun Palvai. "A New Tool for Sharing and Querying of Clinical Documents Modeled Using HL7 Version 3 Standard." *Computer Methods and Programs in Biomedicine, Elsevier* (**CMPB Journal 2013**), Volume 112, Issue 3, December 2013, pp. 529–552. `http://dx.doi.org/10.1016/j.cmpb.2013.07.002`

Vasil Slavov, Praveen Rao, Dinesh Barenkala, and Srivenu Paturi. "Performance of RDF Query Processing on the Intel Single-chip Cloud Computer (SCC)." *Proceedings of the 6th Many-core Applications Research Community Symposium* (**MARC 2012**), Toulouse, France, July 2012, pp. 7–12. `http://hal.archives-ouvertes.fr/hal-00718955`

Vasil Slavov and Praveen Rao. "Towards Internet-Scale Cardinality Estimation of XPath Queries over Distributed XML Data." *Proceedings of the 6th International Workshop on Networking Meets Databases* (**NetDB 2011**), Athens, Greece, June 2011, pp. 1–8. `http://research.microsoft.com/en-us/um/people/srikanth/netdb11/netdb11papers/netdb11-final1.pdf`

Professional Experience

Graduate Research Assistant — **September 2011 – present**
University of Missouri-Kansas City — Kansas City, MO

R&D Developer, Intern — **May 2014 – August 2014**
Bloomberg L.P. — New York, NY

In R&D Transactional Applications, Fixed Income Electronic Trading team
- Developed a regression testing system for multiple large-scale transactional applications in C++ and JavaScript

R&D Developer, Intern — **May 2013 – August 2013**
Bloomberg L.P. — New York, NY

In R&D Transactional Applications, Fixed Income Electronic Trading team
- Developed and evaluated log analytics solutions based on NoSQL systems
- Developed an automated testing system for transactional applications in Python

Network Systems Analyst — **July 2006 – August 2011**
Kansas City Art Institute — Kansas City, MO

Open Systems Analyst — **June 2005 – July 2006**
William Jewell College — Liberty, MO

Languages and Technologies

Languages: C++, Python, C, shell, JavaScript, regex, SPARQL, ANTLR, C#, MPI, LaTeX 2$_\varepsilon$
Cloud Platforms/Distributed Networks: Amazon Elastic Compute Cloud (EC2), IBM SmartCloud Enterprise, PlanetLab
Frameworks/Engines: ElasticSearch, Hadoop, AngularJS, Distributed Hash Tables (Chord)
Tools: git, gdb, Valgrind, CVS
Operating Systems: UNIX/Linux, Mac OS X Server, Windows Server

Technical Experience

Regression testing system (2014). *C++, JavaScript.* Developed a regression testing system for multiple large-scale transactional applications.

RIQ: RDF Indexing on Quadruples (2013 – 2014). *C++, Dablooms, Raptor2, LSH.* Implemented and evaluated a decrease-and-conquer strategy for fast processing of SPARQL queries on very large RDF quadruple datasets (over a billion quads) using Counting Bloom Filters and Locality Sensitive Hashing.

Automated testing system (2013). *Python.* Developed an automated testing suite for replaying production trades of transactional applications.

Log analytics (2013). *AngularJS, ElasticSearch, Kibana, Logstash.* Evaluated log analytics solutions and developed a transaction (Splunk-like) module for Kibana in AngularJS.

XGossip (2010 – 2013). *C++, Sfslite, LSH, Chord DHT, Amazon EC2.* Implemented and evaluated a novel gossip algorithm for estimating the number of XML documents that contain a match for an XPath query in a large-scale network in the presence of churn and failures.

RDF query processing on Intel SCC (2011). *MPICH, C.* Evaluated large-scale RDF query processing on the 48-core Intel Single-chip Cloud Computer (SCC) using task and data parallelism.

SyncMe! (2010). *C#, .NET, Facebook Graph API, Twitter REST API, Google Calendar API, Google OAuth.* In a team, designed and developed a web site for aggregating Twitter and Facebook events on a user's Google Calendar.

Talks

- Presentations
  - Fast Processing of SPARQL Queries on RDF Quadruples. The 17th International Workshop on the Web and Databases (WebDB), Snowbird, UT, June 22, 2014.
  - Tool for Internet-Scale Cardinality Estimation of XPath Queries over Distributed Semistructured Data. The 30th IEEE International Conference on Data Engineering (ICDE), Chicago, IL, April 3, 2014.
  - Performance of RDF Query Processing on the Intel SCC. The 6th Many-core Applications Research Community (MARC) Symposium, Toulouse, France, July 19, 2012.
  - Internet-Scale Cardinality Estimation of XPath Queries over Distributed XML Data. UMKC CSEE Seminar Series, April 2, 2012.
- Guest Lectures
  - Principles of Big Data Management course, UMKC, Fall 2014
  - Data Structures course, UMKC, Fall 2013
  - Introduction to Database Management Systems course, UMKC, Spring 2013 and Fall 2013
  - Advanced Operating Systems course, UMKC, Fall 2010, 2011, and 2013

Service

- Reviewer
  - Journal of Parallel and Distributed Computing, Elsevier (JPDC), 2014
  - International Journal of Distributed Sensor Networks (IJDSN), 2014
- External reviewer
  - International Conference on Database and Expert Systems Applications (DEXA), 2011 - 2014
- Organizer for School of Computing and Engineering E-Week Quiz Contest, UMKC, 2012