

Veselin Raychev

Email: veselin.raychev@gmail.com

Phone: +41 78 867 0124

RESEARCH INTERESTS Program Analysis, Algorithms, Machine Learning

EDUCATION PhD in Computer Science, April 2012 - August 2016
ETH Zurich, Software Reliability Lab
Dissertation: *Learning from Large Codebases*
Advisor: Martin Vechev

M.Sc. in Artificial Intelligence, November 2006 - March 2009
Sofia University, Bulgaria
Thesis: *Automatic Recognition of Positive/Negative Sentiment Polarity in Text*
Advisor: Preslav Nakov

B.Sc. in Informatics, October 2002 - July 2006
Sofia University, Bulgaria

AWARDS AND RECOGNITIONS ACM SIGPLAN Research Highlights on JSNice paper, October 2016
John Atanasoff Award, awarded by the President of Bulgaria, 2016
IBM PhD Fellowship, 2015-2016
Selected to participate in Heidelberg Laureate Forum, 2014
Outstanding Artifact Award, ACM OOPSLA, 2013
SIGPLAN PAC funding award for ACM OOPSLA, 2013
Google OC award for impact on Google Translate, 2008
Silver medal in the Balkan Olympiad in Informatics, Belgrade, Yugoslavia, 2002
Bronze medal in the International Olympiad in Informatics, Korea, 2002
19th place in the ACM ICPC World Finals, San Antonio, TX, 2006
2nd place in the ACM ICPC SEERC, Bucharest, Romania, 2005
29th place in the ACM ICPC World Finals, Shanghai China, 2005
2nd place in the ACM ICPC SEERC, Bucharest, Romania, 2004

TALKS University of Oxford, host: Prof. Hongseok Yang
University of Washington, host: Prof. Emina Torlak
Microsoft Research, Redmond, host: Dr. Madanlal Musuvathi
MIT, host: Prof. Armando Solar-Lezama
UMass Amherst, host: Prof. Emery Berger
IBM T.J. Watson Research, host: Dr. Dave Grove
Columbia University, host: Prof. Junfeng Yang
Yale University, host: Prof. Zhong Shao
Aarhus University, host: Prof. Anders Møller
Dagstuhl seminar, Programming with Big Code, Nov 2015
Dagstuhl seminar, The Future of Refactoring, May 2014

SYSTEMS
DEVELOPED

<http://eventracer.org/>
Race detection for web and android applications. Open source, code was later used also for race detection of SDNs.

<http://jsnice.org/>
JSNice predicts names and type annotations in JavaScript (>150K users).

<http://nice2predict.org/>
Open-source CRF/SSVM system tuned for programming problems.

<http://apk-deguard.com/>
System for performing Android layout deobfuscation.

CONFERENCE
PUBLICATIONS

17. *Probabilistic Model for Code with Decision Trees*
V. Raychev, P. Bielik and M. Vechev
ACM OOPSLA 2016
16. *Statistical Deobfuscation of Android Applications*
B. Bichsel, V. Raychev, P. Tsankov and M. Vechev
ACM CCS 2016
15. *PHOG: Probabilistic Model for Code*
P. Bielik, V. Raychev and M. Vechev
ICML 2016
14. *Learning Programs from Noisy Data*
V. Raychev, P. Bielik, M. Vechev and A. Krause
ACM POPL 2016
13. *Parallelizing user-defined aggregations using symbolic execution*
V. Raychev, M. Musuvathi and T. Mytkowicz
ACM SOSP 2015
12. *Scalable Race Detection for Android Applications*
P. Bielik, V. Raychev and M. Vechev
ACM OOPSLA 2015
11. *Stateless Model Checking of Event-Driven Applications*
C. S. Jensen, A. Moller, V. Raychev, D. Dimitrov and M. Vechev
ACM OOPSLA 2015
10. *Programming with Big Code: Lessons, Techniques and Applications*
P. Bielik, V. Raychev and M. Vechev
SNAPL 2015
9. *Predicting Program Properties from "Big Code"*
V. Raychev, M. Vechev and A. Krause
ACM POPL 2015, Selected for SIGPLAN Research Highlights
8. *Phrase-Based Statistical Translation of Programming Languages*
S. Karaivanov, V. Raychev and M. Vechev
Onward 2014

7. *Code Completion with Statistical Language Models*
V. Raychev, M. Vechev and E. Yahav
ACM PLDI 2014
6. *Commutativity Race Detection*
D. Dimitrov, V. Raychev, M. Vechev and E. Koskinen
ACM PLDI 2014
5. *Effective Race Detection for Event-Driven Programs*
V. Raychev, M. Vechev and M. Sridharan
ACM OOPSLA 2013
4. *Refactoring with Synthesis*
V. Raychev, M. Schaefer, M. Sridharan and M. Vechev
ACM OOPSLA 2013
3. *Automatic Synthesis of Deterministic Concurrency*
V. Raychev, M. Vechev and E. Yahav
SAS 2013
2. *Fast Routing in Very Large Public Transportation Networks Using Transfer Patterns*
H. Bast, E. Carlsson, A. Eigenwillig, R. Geisberger, C. Harrelson, V. Raychev and F. Viger
ESA 2010 (European Symposium on Algorithms)
1. *Language-Independent Sentiment Analysis Using Subjectivity and Positional Information*
V. Raychev and P. Nakov
RANLP 2009

SUPERVISED
STUDENTS

Matteo Panzacchi, M.Sc., Structured Prediction for Programs
Pavol Bielik, M.Sc., EventRacer for Android
Pascal Roos, B.Sc., Effective Statistical Code Completion
Christine Zeller, M.Sc., Statistical Tutoring for MOOCs
Benjamin Bischel, B.Sc., Android De-obfuscation

EMPLOYMENT

Research Intern

Microsoft Research

Redmond, WA, USA

July 2014 - October 2014

Worked in the RiSE group with Madanlal Musuvathi and Todd Mytkowicz on new parallel algorithms useful for large scale logs processing. Internship resulted in an SOSP'15 paper.

Software Engineer

Google Switzerland GmbH

Zurich, Switzerland

November 2006 - March 2012

I worked on the following projects at Google:

Google Maps, October 2008 - March 2012

Designed and implemented key elements of a new state-of-the-art routing engine behind the public transit directions of Google Maps.

Tech Lead of the transit directions team in 2010-2011. Mentored several interns and full-time employees contributing to the project.

Tape backups, November 2006 - September 2008

Designed and implemented key parts of a big scale distributed backup system.

Tech Lead of a team in 2008. Project key result: Gmail was out of beta.

Google Translate, 2007

Trained English to/from Bulgarian systems and developed a morphology component for statistical machine translation to improve translation accuracy.

Software Engineer, Contractor

Motorola Biometrics

Sofia, Bulgaria

November 2005 - October 2006

Implemented a mobile application for fingerprint scanners.

Software Engineering Intern

Google Inc.

Mountain View, CA, USA

Summer 2005

Detected high-level race conditions in the Google File System.

TUTORIALS

Tutorial on "Machine Learning for Code Analytics" at PLDI'15, Portland, OR co-organized with Martin Vechev

WEBSITES

<http://learnbigcode.github.io/>

Includes information about research groups, datasets, challenges and projects learning from large code repositories (also called "Big Code").