

COLLIN McMILLAN

College of William & Mary
McGlothlin-Street 126
Williamsburg, VA 23185

<http://www.cs.wm.edu/~cmc/>
cmc@cs.wm.edu

RESEARCH

Software engineering, source code search, information retrieval in software engineering, repository mining, software evolution and maintenance, traceability, software privacy

EDUCATION

- | | | |
|--------------|---|----------------------|
| Ph.D. | College of William & Mary, Computer Science | <i>expected 2012</i> |
| | Title: <i>Searching, Selecting, and Synthesizing Source Code Components</i> | |
| | Committee: Denys Poshyvanyk, Mark Grechanik, Xipeng Shen, Peter Kemper | |
| M.S. | College of William & Mary, Computer Science | 2009 |
| B.S. | University of Tulsa, Computer Science | 2007 |

HONORS AND AWARDS

- | | |
|--|------|
| Park Graduate Award (\$1,000) | 2012 |
| VSGC Graduate Research Fellowship (\$5,000) | 2011 |
| NSF Travel Grant (\$2,000) | 2011 |
| ACM SigSoft Travel Grants (2 x \$500) | 2011 |
| Microsoft Student Research Competition (\$500) | 2011 |

PUBLICATIONS

Journal Papers

McMillan C., Grechanik M., Poshyvanyk D., Fu. C., and Xie Q., “Exemplar: A Source Code Search Engine for Finding Highly-Relevant Applications,” *IEEE Transactions on Software Engineering (TSE)*, to appear

Conference Papers

McMillan C., Hariri N., Poshyvanyk D., Cleland-Huang J., and Mobasher B., “Recommending Source Code for use in Rapid Software Prototypes,” in *Proc. of 34th ACM/IEEE International Conference on Software Engineering (ICSE’12)*, Zurich, Switzerland, June 2-9, 2012, to appear (21% acceptance rate)

McMillan C., Grechanik M., Poshyvanyk D., “Detecting Similar Software Applications,” in *Proc. of 34th ACM/IEEE International Conference on Software Engineering (ICSE’12)*, Zurich, Switzerland, June 2-9, 2012, to appear (21% acceptance rate)

McMillan C., Linares-Vásquez M., Poshyvanyk D., and Grechanik M., “Categorizing Software Applications for Maintenance,” in *Proc. of 27th International Conference on Software Maintenance (ICSM’11)*, Williamsburg, VA, September 25-30, 2011, pp. 343-352 (28% acceptance rate)

McMillan C., Grechanik M., Poshyvanyk D., Xie Q., and Fu C., “Portfolio: Finding Relevant Functions and Their Usages,” in *Proc. of 33rd IEEE/ACM International Conference on Software Engineering (ICSE’11)*, Honolulu, HI, USA, May 21-28, 2011, pp 111-120 (14% acceptance rate)

McMillan C., Grechanik M., Poshyvanyk D., Xie Q., and Fu C., “Portfolio: A Search Engine for Finding Functions and Their Usages,” in *Proc. of 33rd IEEE/ACM International Conference on Software Engineering (ICSE’11)*, Formal Research Tool Demonstration, Honolulu, Hawaii, USA, May 21-28 2011, pp. 1043-1045 (36% acceptance rate)

Grechanik M., **McMillan C.**, DeFerrari L., Comi M., Crespi S., Poshyvanyk D., Fu C., Xie Q., and Ghezzi C., “An Empirical Investigation into a Large-Scale Java Open Source Code Repository,” in *Proc. of 4th International Symposium on Empirical Software Engineering and Measurement (ESEM’10)*, Bolzano-Bozen, Italy, September 16-17, 2010, pp. 1-11 (29% acceptance rate)

Grechanik M., Fu C., Xie Q., **McMillan C.**, Poshyvanyk D., and Cumby C., “A Search Engine For Finding Highly-Relevant Applications,” in *Proc. of 32nd ACM/IEEE International Conference on Software Engineering (ICSE’10)*, Cape Town, South Africa, May 2-8, 2010, pp. 475-484 (13.6% acceptance rate)

Grechanik M., Fu C., Xie Q., **McMillan C.**, Poshyvanyk D., and Cumby C., “Exemplar: EXECutable exAMPLeS ARchive,” in *Proc. of 32nd ACM/IEEE International Conference on Software Engineering (ICSE’10)*, Formal Research Tool Demonstration, Cape Town, South Africa, May 2-8, 2010, pp. 259-262 (25% acceptance rate)

Workshop Papers

McMillan C., “Searching, Selecting, and Synthesizing Source Code,” in *Proc. of 33rd IEEE/ACM International Conference on Software Engineering (ICSE’11)*, Doctoral Symposium, Honolulu, Hawaii, May 21-28 2011, pp. 1124-1125

McMillan C., “Finding Relevant Functions in Millions of Lines of Code,” in *Proc. of 33rd IEEE/ACM International Conference on Software Engineering (ICSE’11)*, ACM Student Research Competition, Honolulu, Hawaii, USA, May 21-28 2011, pp. 1170-1172

McMillan C., Poshyvanyk, D., and Grechanik, M., “Recommending Source Code Examples via API Call Usages and Documentation,” in *Proc. of 2nd International Workshop on Recommendation Systems for Software Engineering (RSSE’10)*, Cape Town, South Africa, May 4, 2010, pp 21-25

Louthan G., **McMillan C.**, Johnson C., Hale J., “Toward Robust and Extensible Automatic Protocol Identification,” in *Proc. of 10th International Conference on Internet Computing (ICOMP09)*, Las Vegas, USA, July 13-16, 2009, pp. 104-108

McMillan C., Poshyvanyk D., Revelle M., “Combining Textual and Structural Analysis of Software Artifacts for Traceability Link Recovery,” in *Proc. of 4th International Workshop on Traceability in Emerging Forms of Software Engineering (TEFSE’09)*, Vancouver, British Columbia, Canada, May 18, 2009, pp. 41-48

Journal Papers under Review

McMillan C., Grechanik M., Poshyvanyk D., Xie Q., Fu C., “Portfolio: Searching for Relevant Functions and Their Usages in Millions of Lines of Code,” *ACM Transactions on Software Engineering and Methodology (TOSEM)*, under major revision

Linares-Vásquez M., **McMillan C.**, Poshyvanyk D., Grechanik M., “Automatically Classifying Source and Byte Code into Domain Categories,” *Empirical Software Engineering (ESE)*

PRESENTATIONS AND INVITED LECTURES

Invited Talk, “Searching, Selecting, and Synthesizing Source Code Components,” Wayne State University, November 15, 2011.

Conference Paper Presentation, “Categorizing Software Applications for Maintenance,” IEEE International Conference on Software Maintenance (ICSM), September 29, 2011.

Conference Paper Presentation, “Portfolio: Finding Relevant Functions and Their Usages,” ACM/IEEE International Conference on Software Engineering (ICSE), May 25, 2011.

Formal Tool Demonstration, “Finding Relevant Functions in Millions of Lines of Code,” ACM/IEEE International Conference on Software Engineering (ICSE), May 26, 2011.

Poster Presentation, “Searching, Selecting, and Synthesizing Source Code,” ACM IEEE International Conference on Software Engineering (ICSE), May 24, 2011.

Invited Lecture, “Design Patterns in Software Engineering,” Software Engineering Class at William & Mary, March 14, 2011.

Formal Tool Demonstration, “Exemplar: EXEcutable exaMPLES ARchive,” ACM/IEEE International Conference on Software Engineering (ICSE), May 5, 2010.

Workshop Paper Presentation, “Combining Textual and Structural Analysis of Software Artifacts for Traceability Link Recovery,” Workshop on Traceability in Emerging Forms of Software Engineering (TEFSE), May 18, 2009.

STUDENT SUPERVISION

Mario Linares-Vásquez, *Ph.D. Student*, advisor Denys Poshyvanyk, W&M (2010)

I supervised Mario during our work towards our paper at ICSM'11 on categorizing software applications. I oversaw Mario's implementation, evaluation, and writing efforts in a distance-learning program while Mario was studying in Bogotá, Columbia from July 2010 to August 2011. I continue to supervise Mario during his six-month residency at William & Mary starting August 2011.

Tathagata Dasgupta, *M.S. Student*, advisor Mark Grechanik, UIC (2011)

I supervised Tathagata's implementation of RESIST, a framework for redaction of private information from source code, which resulted in a submission to ICSE'12 (under review). I worked closely with Tathagata starting May 2011, including collaborative programming and critiquing the project design. I continue to oversee his work towards a M.S. Thesis until May 2012.

Luca DiMinervino, *M.S. Student*, advisor Mark Grechanik, UIC (2009)

I supervised Luca's M.S. Thesis during his implementation of key components of the source code search engine Portfolio. I managed a team of nine M.S. students who worked with Luca at the University of Illinois at Chicago. I designed programming projects, and coordinated the efforts of the students including Luca. This effort culminated in a publication at ICSE'11 and a journal submission to TOSEM (under review).

Evan Moritz, *Undergraduate*, advisor Denys Poshyvanyk, W&M (2011)

I worked closely with Evan to build a visualization for the API-learning tool ExPort during the Spring of 2011. I integrated his class assignments with an ongoing research project, and Evan became a key contributor to the project.

PROFESSIONAL EXPERIENCE

Graduate Research Assistant, advisor Denys Poshyvanyk, W&M (2009-2012)

I participated in several research projects funded by the NSF as a member of the SEMERU Lab. I developed and maintained the source code search engines Exemplar and Portfolio, as well as the software privacy framework RESIST, in collaboration with Accenture Technology Labs and the University of Illinois at Chicago. I worked with multiple Information Retrieval, data mining, and program analysis techniques.

Technician, supervisor Phil Kearns, W&M (2007-2009)

I built and maintained the William & Mary Computer Science department's computer network as part of a team of four technicians. I held a position of trust, including administrator access to all machines. My position was very programming intensive and required detailed systems knowledge of Linux, network security, and maintenance of legacy software applications.

Undergraduate Research Assistant, advisor John Hale, TU (2004-2007)

I implemented several key components for NSF-funded projects with the ESG Lab. I designed and wrote the network analysis tool SAND, and was responsible for data collection tools supporting the network security framework NOVA.

PROFESSIONAL ACTIVITIES

Organizing Committee

2011 Student Volunteer Chair

27th IEEE International Conference on Software Maintenance

11th IEEE International Work. Conf. on Source Code Analysis and Manipulation

13th IEEE International Symposium on Web Systems Evolution

6th IEEE Inter. Work. on Visualizing Software for Understanding and Analysis

Program Committee

2012 20th ACM Inter. Symp. on Foundations of Software Engineering, Tool Demo Track

2011 27th IEEE International Conference on Software Maintenance, Tool Demo Track

Community Service

2011 Treasurer, W&M Computer Science Graduate Student Association

2010 Student Volunteer, 32nd ACM/IEEE Inter. Conference on Software Engineering

External Reviewer for:

2011 27th IEEE International Conference on Software Maintenance

19th IEEE International Conference on Program Comprehension, Industry

12th International Workshop on the Principles of Software Evolution

Journal of Software Maintenance and Evolution: Research and Practice

IEEE Transactions on Software Engineering

8th ACM Symposium on the Foundations of Software Engineering, Tool Demos

2010 26th IEEE International Conference on Software Maintenance

18th IEEE International Conference on Program Comprehension

14th European Conference on Software Maintenance and Reengineering

Journal of Software Maintenance and Evolution: Research and Practice

Software Quality Journal

Empirical Software Engineering: An International Journal

2009 25th IEEE International Conference on Software Maintenance

OTHER

FAA Student Pilot Certificate, Airplane, Single-Engine Land

United States Citizen

REFERENCES

Dr. Denys Poshyvanyk

Assistant Professor, College of William & Mary, Williamsburg, VA

Dr. Mark Grechanik

Senior Research Scientist, Accenture Technology Labs, Chicago, IL

Dr. Jane Cleland-Huang

Associate Professor, DePaul University, Chicago, IL

Dr. Andrian Marcus

Associate Professor, Wayne State University, Detroit, MI

Dr. John Hale

Professor, The University of Tulsa, Tulsa, OK