KANG HONG JIN

Ph.D Student
School of Information Systems
Singapore Management University

hjkang.2018@phdcs.smu.edu.sg
kanghj.github.io

EDUCATION

Ph.D
2019 -
School of Information Systems
Singapore Management University
GPA: 3.91/4

BACHELOR OF COMPUTING
2012-2016
School of Computing
National University of Singapore
First Class Honours/Highest Distinction

RESEARCH INTERESTS

- Specification Mining
- Machine Learning on Software Engineering
- Test Case Generation

PUBLICATIONS

Active Learning of Discriminative Subgraph Patterns for API Misuse Detection
by **Hong Jin KANG**; David LO.
IEEE Transactions on Software Engineering
https://ieeexplore.ieee.org/document/9392340/

Adversarial Specification Mining
by **Hong Jin KANG**; David LO.
ACM Transactions on Software Engineering and Methodology, Vol. 30, No. 2
https://arxiv.org/abs/2103.15350

IoTBox: Sandbox Mining to Prevent Interaction Threats in IoT Systems
by **Hong Jin KANG**; David LO; Sheng Qin SIM
IEEE International Conference on Software Testing, Verification and Validation (ICST 2021)
https://kanghj.github.io/publications/iot-sandbox-camera-ready.pdf

AndroEvolve: Automated Update for Android Deprecated-API Usages
by Stefanus Agus HARYONO; Ferdian THUNG; David LO; Lingxiao JIANG; Julia LAWALL; **Hong Jin KANG;** Lucas SERRANO; Gilles MULLER
ACM/IEEE International Conference on Software Engineering 2021 (ICSE 2021) – Demonstrations
https://arxiv.org/abs/2012.07259

BugsInPy: a database of existing bugs in Python programs to enable controlled testing and debugging studies
by Ratnadira WIDYASARI; Sheng Qin SIM; Camellia LOK; Haodi QI; Jack PLAN; Qijin TAY; Constance TAN; Fiona WEE; Jodie Ethelda TAN; Yuheng YIEH; Brian GOH; Ferdian Thung; **Hong Jin KANG**; Thong HOANG; David LO; Eng Lieh OUH
ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2020)
https://ink.library.smu.edu.sg/sis_research/5630/

CC2Vec: Distributed representations of code changes,
by Thong HOANG; **Hong Jin KANG**; Julia LAWALL; David LO.
ACM/IEEE International Conference on Software Engineering 2020 (ICSE 2020)
https://arxiv.org/abs/2003.05620

Automatic Android Deprecated-API Usage Update by Learning from Single Updated Example
by Stefanus Agus HARYONO; Ferdian THUNG; **Hong Jin KANG**; Lucas SERRANO; Gilles MULLER; Julia LAWALL; David LO; Lingxiao JIANG.
International Conference on Program Comprehension 2020 - Early Research Achievements Track. 2020 (ICPC ERA 2020)
https://arxiv.org/abs/2005.13220

Towards generating transformation rules without examples for Android API replacement
by Ferdian THUNG; **Hong Jin KANG**; Lingxiao JIANG; David LO.
IEEE International Conference on Software Maintenance and Evolution 2019 (ICSME 2019)
https://ink.library.smu.edu.sg/sis_research/4824/

Assessing the generalizability of code2vec token embeddings
by **Hong Jin KANG**; Tegawende F BISSYANDE.; David LO.
ACM/IEEE International Conference on Automated Software Engineering 2019 (ASE 2019)
http://www.mysmu.edu/faculty/davidlo/papers/ase19-code2vec.pdf

Semantic patches for Java program transformation
by **Hong Jin KANG**; Ferdian THUNG; Julia LAWALL; Gilles MULLER; Lingxiao JIANG; David LO.
European Conference on Object-Oriented Programming 2019 (ECOOP 2019)
https://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=5488&context=sis_research

A Comparison of Word Embeddings for English and Cross-Lingual Chinese Word Sense Disambiguation
By **Hong Jin KANG;** Tao CHEN; Muthu Kumar CHANDRASEKARAN; Min-Yen KAN.
The COLING Workshop on Natural Language Processing Techniques for Educational Applications Workshop (NLP-TEA-3 2016)
https://arxiv.org/abs/1611.02956

ONGOING WORK

AndroEvolve: Automated Android API Update with Data Flow Analysis and Variable Denormalization
by Stefanus Agus HARYONO; Ferdian THUNG; David LO; Lingxiao JIANG; Julia LAWALL; **Hong Jin KANG;** Lucas SERRANO; Gilles MULLER
Under Major Revision at the journal, Empirical Software Engineering

BiasFinder: Metamorphic Test Generation to Uncover Bias for Sentiment Analysis Systems.
Muhammad Hilmi ASYROFI; Imam Nur Bani YUSUF; **Hong Jin KANG;** Ferdian THUNG; Zhou YANG; David LO:
Under Major Revision at the journal, Transactions on Software Engineering

Test generation for uncovering vulnerabilities from third-party libraries
Hong Jin KANG; Truong Giang NGUYEN; Bach LE; Corina PASAREANU; David LO;
Currently under review at the conference, Automated Software Engineering

WORKING EXPERIENCE

AUG 2018 – DEC 2018
RESEARCH ENGINEER, SMU

- Helped to develop Coccinelle4J, a port of Coccinelle for the Java programming language. Coccinelle is a widely adopted program matching and transformation tool for C systems software, including the Linux kernel. Over 6000 commits have been in the Linux kernel with the help of Coccinelle in the past 10 years, and we apply it to the migration of deprecated Android API usages.

2018
BACKEND ENGINEER, GRAB

- Worked with cutting-edge tools and managed cloud-based infrastructure.
- Owned a microservice on the critical path of all bookings. Improved testing to identify single points of failure, improving the resilience of the service to ensure uptime and SLAs are met.

2016 –2018
SOFTWARE ENGINEER, WORKS APPLICATIONS CO. LTD

- Worked on enterprise-grade HR Enterprise Resource Management software.
- Helped mentored several junior engineers.
- Full-stack development from the back-end using Spring, management of Cassandra, Elasticsearch, to implementing a new user experience on the front-end with modern frameworks.

AUG 2015 – MAY 2016
STUDENT SOFTWARE ENGINEER, TEAMMATES @ NUS

- Involved in both planning and development of the project. Planned and managed weekly releases.
- Part of the team that introduced static analysis tools and CI to the project.

MAY 2014 – AUG 2014
INTERN, TEAMIE

- Developed several new features, including a feature that summarizes users' activity logs to identify interesting parts of their history on the platform.

SKILLS

- General Software Engineering (Git, various IDEs, bash, python, Java)
- Functional Programming (OCaml, Scala)
- Machine learning and Deep Learning Frameworks (scikit-learn, tensorflow)
- Web Programming (Javascript, Spring framework)
- Databases (MySQL, Cassandra, Elasticsearch)
- Declarative programming (answer set programming, clingo)