HOANG VAN DUC THONG (JAMES)

(+84) 907 410 960 ⋄ vdthoang.2016@phdcs.smu.edu.sg

Homepage: https://hvdthong.github.io/

EDUCATION

Singapore Management University, Singapore *Aug. 2016 - Aug. 2020*
School of Computing and Information Systems
Ph.D. in Computer Science

Konkuk University, South Korea *Mar. 2009 - Jun. 2011*
Department of Advanced Technology Fusion
Master of Computer Science

Ho Chi Minh City University of Technology, Vietnam *Sep. 2004 - Feb. 2009*
Department of Computer Science
Bachelor of Computer Science

RESEARCH INTEREST

- **Software repository mining**: Code understanding and inference; bug localization; bug fixing patches; just-in-time defect prediction

- **Social network analysis and mining:** Tracking topics in document streams; modelling users and communities

WORKING EXPERIENCE

Jr. Data Scientist; Trusting Social; Vietnam *From Apr. 2021 - Now*

- Building a model to evaluate credit scores for financial access

Research Fellow; Living Analytics Research Center; Singapore *From Sep. 2020 - Mar. 2021*

- Building a recommendation system for students' courses

Research Intern; Fujitsu Lab of America; Silicon Valley - USA *Sep. 2019 - Dec. 2019*

- Developing a system for checking the quality of AI models

Research Engineer; Living Analytics Research Center; Singapore *May 2015 - Jul. 2016*

- Developing a dashboard for tracking transportation systems in Singapore using social media

- Developing a system for detecting transportation events and perform sentiment analysis

Research Engineer; Ghent University; Belgium *Feb. 2012 - Apr. 2015*

- Developing a ranking system for finding relevant documents given a query

- Employing a budget aware learning, by exploiting cheap approximate annotation, to minimize human annotation effort

Research Assistant; Konkuk University; South Korea *Apr. 2009 - May 2011*

- Finding powerful users in social network

- Identifying a relationship among users in social media

TEACHING

Singapore Management University *Aug. 2016 - Aug. 2020*

- Business and data management, as teaching assistant (IS110)

- Computation thinking, as teaching assistant (IS103)

SKILLS

Programming Languages	**Python**, **Java**, Matlab, R, C/C++/C#
Deep Learning Frameworks	**Pytorch**, **Tensorflow**, **Keras**, Theano
Machine Learning Libraries	**Sklearn**, **Numpy**, **Pandas**, Weka
Operating Systems	**Linux**, **Windows**
Database Management	**MS SQLServer**, **MySQL**, ElasticSearch
Editor	**Latex**, **MS Office**

PUBLICATION

[12] Xu Bowen, Thong Hoang, Abhishek Sharma, Xin Xia, David Lo; **Post2Vec: Learning Distributed Representations of Stack Overflow Posts;** IEEE Transactions on Software Engineering (TSE), 2021.

[11] Ratnadira Widyasari, Sheng Qin Sim, Camellia Lok, Haodi Qi, Jack Phan, Qijin Tay, Constance Tan, Fiona Wee, Jodie Ethelda Tan, Yuheng Yieh, Brian Goh, Ferdian Thung, Hong Jin Kang, Thong Hoang, David Lo, and Eng Lieh Ouh; **BugsInPy: A database of existing bugs in Python programs to enable controlled testing and debugging studies;** The Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (FSE), Short Paper, Sacramento, United States, 2020.

[10] Roy Ka-Wei Lee, Thong Hoang, Richard J. Oentaryo, David Lo; **Keen2Act: Activity Recommendation in Online Social Collaborative Platforms;** International Conference on User Modeling, Adaptation and Personalization, Short Paper, Genoa, Italy, 2020.

[9] Thong Hoang, Hong Jin Kang, Julia Lawall, David Lo; **CC2Vec: Distributed Representations of Code Changes;** The 42nd International Conference on Software Engineering (ICSE), Seoul, South Korea, 2020.

[8] Thong Hoang, Julia Lawall, Richard J Oentaryo, Yuan Tian, David Lo; **PatchNet: Hierarchical Deep Learning-Based Stable Patch Identification for the Linux Kernel;** IEEE Transactions on Software Engineering (TSE), 2019.

[7] Thong Hoang, Hoa Khanh Dam, Yasutaka Kamei, David Lo, Naoyasu Ubayashi; **DeepJIT: An End-To-End Deep Learning Framework for Just-In-Time Defect Prediction;** The 16th Mining Software Repositories Conference (MSR 2019), Montreal, Canada, May 2019.

[6] Thong Hoang, Julia Lawall, Richard J Oentaryo, Yuan Tian, David Lo; **PatchNet: A Tool for Deep Patch Classification;** The 41st ACM/IEEE International Conference on Software Engineering (ICSE 2019), Demonstrations Track, Montreal, Canada, May 2019.

[5] Thong Hoang, Richard J Oentaryo, Tien-Duy Bui Le, David Lo; **Network-clustered multi-modal bug localization;** IEEE Transactions on Software Engineering (TSE), 2018.

[4] Thong Hoang, Pei Hua Cher, Philips Prasetyo Kokoh, Ee-Peng Lim; **Crowdsensing and Analyzing Micro-Event Tweets for Public Transportation Insights;** The 2nd International workshop on Big Data for Sustainable Development with IEEE Big Data (IEEE BigData), Washington, USA, December 2016.

[3] Thong Hoang, Thomas Demeester, Johannes Deleu, Piet Demeester, Chris Develder; **Ugent Participation in the Microblog Track 2012;** The 22nd Text Retrieval Conference (TREC 2012), Gaithersburg, USA, October 2012.

[2] Thong Hoang, Thomas Demeester, Chris Develder, Hyoseop Shin; **Effectiveness of Learning to Rank for finding User Similarity in Social Media;** The 12th Dutch-Belgian Information Retrieval Workshop (DIR 2012), Ghent, Belgium, April 2012.

[1] Nhat Hai Phan, Thong Hoang, and Hyoseop Shin; **Adaptive Combination of Tag and Link-based User Similarity in Flickr;** The 18th ACM International Conference on Multimedia (ACM MM 2010), Firenze, Italy, October 2010.

AWARDS

2020: Placed on the Dean's List of SMU Postgraduate Program

2019: Awarded SMU Presidential Doctoral Fellowship

2016: Full Ph.D. scholarship from Ministry of Education, Singapore

SCIENTIFIC COMMUNITY ACTIVITIES

Journal Review: Neurocomputing (2018, 2019, 2020), IST (2019, 2020), DAMI (2019), SCP (2020), JSS (2020), IEEE TSE (2020), SNCS (2020)
Conference Review (and Sub-Review): SCAM (2019), ICECCS (2019), SANER (2020), ICSME (2020), KDD (2020)

Student Volunteer:

- The 18th IEEE International Conference on Data Mining (ICDM 2018)

- The 31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016)

GRANTS

- Assisted in writing part of funding proposal: *Uncovering Vulnerabilities in Machine Learning Frameworks via Software Composition Analysis and Directed Grammar Based Fuzzing*, a joint proposal between Singapore Management University, Veracode, Nanjing University, Carnegie Mellon University, NASA, and University of Melbourne University. The funding was accepted by National Research Foundation (Singapore), 2020.

- Assisted in writing part of funding proposal: *Making Big Code Active: From Billions of Code Tokens to Automation*, a joint proposal between Singapore Management University and Zhejiang University. The funding was accepted by Singapore Data Science Consortium, 2020.

REFERENCES

Dr. David Lo
Full Professor
Singapore Management University
Email: davidlo@smu.edu.sg

Dr. Hoa Khanh Dam
Associate Professor
University of Wollongong
Email: hoa@uow.edu.au

Dr. Julia Lawall
Inria Senior Research Scientist
Sorbonne University, Inria, LIP6
Email: Julia.Lawall@lip6.fr

Dr. Richard Oentaryo
Principal Jr. Data Scientist
QuantumBlack, McKinsey & Company
Email: richard_oentaryo@mckinsey.com