

DUC BUI

(734) 486-8970
ducbui@umich.edu · ducbui.com

My research lies at the intersection of **user privacy**, **program analysis** and **natural language processing**. I am building end-to-end automated systems and formal models that analyze the consistency between privacy policies and actual data collection of mobile and web apps. I previously worked on **mobile systems and networking** where I developed novel quality-of-service-guaranteed energy-efficient technologies for the Chrome web browser and real-time video streaming.

EDUCATION

University of Michigan , Ann Arbor, MI	2017 – present
<ul style="list-style-type: none">• PhD Candidate in Computer Science · Advisor: Professor Kang G. Shin	
Korea Advanced Institute of Science and Technology (KAIST) , South Korea	2011 – 2013
<ul style="list-style-type: none">• MSc in Computer Science · GPA 4.0/4.3 (96.7/100) · Outstanding Master's Thesis Award	
Hanoi University of Science and Technology , Vietnam	2005 – 2010
<ul style="list-style-type: none">• BSc in Computer Science · GPA 8.5/10 (top 1% of CS Dept.)	

INDUSTRIAL EXPERIENCE

Facebook , Menlo Park, CA	Research Intern	May – Aug 2019
<ul style="list-style-type: none">• <i>Personalized Federated Learning</i>: Designed the 1st privacy-preserving user representation learning for Federated Learning. My techniques improve 8-51% AUC/accuracy than non-personalized models and provide similar performance with centralized approaches while preserving user privacy. Facebook R&D continued to use my implementation (3000+ lines of code) after my internship. [PyTorch, PyText, Python]• This work was submitted to ICLR 2020 and published on arXiv.		
Google , Mountain View, CA	Software Engineering Intern	May – Aug 2017
<ul style="list-style-type: none">• <i>Performance metrics for the Chrome web browser</i>: Added congestion tracking metrics to Chrome to quantify page load times and the responsiveness of foreground tabs when the browser loads multiple tabs simultaneously. The metrics are deployed in production. Chrome engineers use them to direct optimization efforts to improve the user experience when loading many tabs. I made 16 code changes (added 3000+ and deleted 500+ lines of code) to the Chrome codebase. [C++]• <i>Power consumption estimation from CPU usage</i>: Investigated the correlation between Chrome's CPU usage and power consumption to estimate the power profiles of devices in the wild. [Python, JavaScript]		
Samsung Electronics , Suwon, South Korea	Research Intern	May – Sep 2015
<ul style="list-style-type: none">• <i>Transparent cross-device resource sharing</i>: Used real-time streaming protocols (RTSP/RTCP/RTP) to improve camera preview frame rate by 6X and reduce photo capture time across two smartphones by 4X, compared to the state of the art (Rio, MobiSys 2014); developed Unified Resource Management Framework with 11,000+ lines of C/C++ code that manages resources (camera, sensors, and apps) across multiple heterogeneous-platform mobile devices. [C, C++, CMake, Tizen OS]		
Microsoft Research , Beijing, China	Research Intern	Jan – Jun 2014
<ul style="list-style-type: none">• <i>Energy-efficient mobile web browsing</i>: Reduced the whole-system energy consumption of the Google Chrome mobile web browser by 24.4% with no perceivable impact on page load time; analyzed inefficiencies, developed, and implemented energy-saving techniques for the Chrome and Firefox web browsers on Android. Compared with state-of-the-art solutions, my techniques leverage browser internals (process/thread structure and resource fetching/processing pipelines) and the heterogeneous multi-core big.LITTLE CPU architecture rather than based on characteristics of web pages. [C++, Python, .NET]• This work was published in a top-tier conference (ACM MobiCom) and highlighted on an ACM SIGMOBILE magazine and online news (Microsoft Research Blog, Phys.org, and On MSFT); open source on GitHub: energy_efficient_web_page_loading_browser_profiler.		

ACADEMIC EXPERIENCE

University of Michigan , Ann Arbor, MI	Research Assistant	Sep 2017 – present
<ul style="list-style-type: none">• <i>Assessment of privacy risks in mobile and web applications/services (PhD thesis)</i>: I developed 5 End-to-End (E2E) systems to automatically assess the privacy risks of mobile and web apps through the analysis of privacy policies, app execution and user interfaces. Of these, two systems analyze the flow-to-policy consistency between the privacy statements in privacy policies and the data flows of mobile/web apps. Two other systems detect the inconsistencies between the opt-out settings and their enforcement of websites and online trackers. Finally, a system analyzes website privacy policies and provides an easy-to-understand presentation to help users beware of the practices performed on their data. [Dynamic analysis of Android apps and web apps, natural language processing (NLP), formal modeling, user study]• This work results in 5 papers: two were published and three are under review of top-tier conferences (e.g., ACM CCS).		

KAIST, Daejeon, South Korea

Research Assistant

Sep 2011 – Jun 2017

- *Energy-efficient multi-link real-time streaming*: I created the first LTE-enabled prototype implementation that aggregates bandwidth over multiple asymmetric mobile wireless interfaces (WiFi & LTE) for energy-efficient real-time delivery; provided quality-of-service (QoS) for high-bitrate video streaming while reducing energy consumption by leveraging Radio Resource Control (RCC) states of LTE. Compared to state-of-the-art solutions, my method does not require any support on existing servers and changes to the existing Internet infrastructure. [C, Java, Android]
- This work was published in a top-tier conference (IEEE RTSS); open source on GitHub: [greenbag](#).

PUBLICATIONS

1. *Consistency Analysis of Data-Usage Purposes in Mobile Apps*.
Duc Bui, Yuan Yao, Kang G. Shin, Jong-Min Choi, and Junbum Shin
ACM SIGSAC Conference on Computer and Communications Security (**CCS**), Virtual event, South Korea, November 2021.
2. *Automated Extraction and Presentation of Data Practices in Privacy Policies*.
Duc Bui, Kang G. Shin, Jong-Min Choi, and Junbum Shin
Privacy Enhancing Technologies Symposium (**PETS**), Virtual conference, July 2021.
3. *Federated User Representation Learning*.
Duc Bui, Kshitiz Malik, Jack Goetz, Honglei Liu, Seungwhan Moon, Anuj Kumar, and Kang G. Shin.
arXiv preprint arXiv:1909.12535, 2019.
4. *Active Federated Learning*.
Jack Goetz, Kshitiz Malik, **Duc Bui**, Seungwhan Moon, Honglei Liu, and Anuj Kumar.
Workshop on Federated Learning for Data Privacy and Confidentiality (in Conjunction with **NeurIPS**) 2019.
5. *Cross-Platform Support for Rapid Development of Mobile Acoustic Sensing Applications*.
Yu-Chih Tung, **Duc Bui**, and Kang G. Shin.
ACM International Conference on Mobile Systems, Applications and Services (**MobiSys**) 2018. (**27%** acceptance rate)
6. *Mobile Plus: Mobile platform for Transparent Sharing of Functionalities Across Devices*.
Sangeun Oh, Hyuck Yoo, Dae R. Jeong, **Duc Bui**, and Insik Shin.
ACM International Conference on Mobile Systems, Applications and Services (**MobiSys**) 2017. (**18%** acceptance rate)
7. *Demo: Mobile Plus: Mobile platform for Transparent Sharing of Functionalities Across Devices*.
Sangeun Oh, Hyuck Yoo, Daelyong Jeong, Sooyoung Park, **Duc Bui**, Sungsoo Moon, and Insik Shin.
ACM International Conference on Mobile Systems, Applications and Services (**MobiSys**) 2016.
8. *Rethinking Energy-Performance Trade-Off in Mobile Web Page Loading*.
Duc Bui, Yunxin Liu, Hyosu Kim, Insik Shin, and Feng Zhao.
ACM SIGMOBILE GetMobile Magazine (**Research highlights**), April 2016.
9. *Rethinking Energy-Performance Trade-Off in Mobile Web Page Loading*.
Duc Bui, Yunxin Liu, Hyosu Kim, Insik Shin, and Feng Zhao.
ACM International Conference on Mobile Computing and Networking (**MobiCom**) 2015. (**18%** acceptance rate)
10. *GreenBag: Energy-efficient Bandwidth Aggregation for Real-time Streaming in Heterogeneous Mobile Wireless Networks*.
Duc Bui, Kilho Lee, Sangeun Oh, Hyojeong Shin, Insik Shin, Honguk Woo, and Daehyun Ban.
IEEE Real-Time Systems Symposium (**RTSS**) 2013. (**22%** acceptance rate)
11. *A Case Study of the Application of Dynamic Symbolic Execution to Real-World Binary Programs*.
Duc Bui, Yunho Kim, and Moonzoo Kim.
Korea Conference on Software Engineering (**KCSE**) 2012.
12. *A method of verifying web service composition*.
Thang Huynh, Quynh Pham, and **Duc Bui**.
ACM International Symposium on Information and Communication Technology (**SoICT**) 2010.

TALKS

- *VinAI research lab seminar series*, October 2019: Federated User Representation Learning.
- *MobiCom 2015 oral presentation*, September 2015: Rethinking Energy-Performance Trade-Off in Mobile Web Page Loading.
- *RTSS 2013 oral presentation*, December 2013: GreenBag: Energy-efficient Bandwidth Aggregation for Real-time Streaming in Heterogeneous Mobile Wireless Networks.
- *KCSE 2012 oral presentation*, February 2012: A Case Study of the Application of Dynamic Symbolic Execution to Real-World Binary Programs.
- *Workshop on Software Analysis for Error-free Computing in Hong Kong University of Science and Technology (HKUST)*, January 2012: A Case Study of the Application of Dynamic Symbolic Execution to Real-World Binary Programs.

PROFESSIONAL SERVICES

- *External review*: ACM Wireless of the Students, by the Students, and for the Students (S3) Workshop.

HONORS AND AWARDS

- | | |
|------|--|
| 2016 | Naver PhD Fellowship by Naver Corp. (the largest search engine in South Korea) for excellent PhD students. |
| 2015 | Qualcomm Innovation Award finalist, being in the top ten out of 37 final-round projects. |
| 2015 | Microsoft Research Asia Fellowship Nomination Award , being one of 90 students selected from top Asia universities. |
| 2014 | Microsoft Research Asia Excellent Award in the Stars of Tomorrow Internship Program. |
| 2014 | Outstanding Master's Thesis Award by Computer Science Department, KAIST. |
| 2010 | Korean Government Scholarship for the master's program at KAIST. |
| 2009 | Vietnam Ministry of Information and Communications Scholarship for outstanding students. |

REFERENCES

(Available upon request)