Kangkook Jee

NEC Laboratories America
400 Independence Way Suite 200
Princeton, NJ
https://kangkookjee.github.io

Phone: (609) 951-2909
Email: kjee@nec-labs.com
Alt: kangkook.jee@gmail.com

PARTICULARS

EDUCATION

Columbia University in the city of New York — New York, NY
Ph. D. in Computer Science — *2008-2014*

Columbia University in the city of New York — New York, NY
M. Phil. in Computer Science — *Jun 2012*

Columbia University in the city of New York — New York, NY
M. Sc. in Computer Science — *Dec 2007*
Distinction in Research

Korea University — Seoul, South Korea
B.Sc. majoring Mathematics minoring Computer Science — *Feb 2000*

CURRENT STATUS

U.S. Permanent Resident, Citizen of South Korea.

DISSERTATION

Title: "On Efficiency and Accuracy of Data Flow Tracking Systems"
Advisor: Prof. Angelos D. Keromytis

My thesis develops a Data Flow Tracking (DFT) framework that works with commodity software and hardware. Also, I designed a set of techniques to reduce the slowdown of DFT frameworks firstly by combining static and dynamic analysis, and by decoupling analysis from execution and using spare CPU cores to run them in parallel.

RESEARCH INTERESTS

My research interests span the areas of system and software security. I have specific interests in software instrumentation leveraging hypervisor and compiler technologies, efficient system event auditing, cyber threat intelligence, and binary analysis.

WORK EXPERIENCE

- **NEC Laboratories America** Sep 2014 - Present.

  Researcher, Computer Security Department: Designed and implemented end-point security solutions for enterprise computers, IoT devices, Automobiles, and ICS devices. The platforms are designed to collect and process that large-scale data-set to support ML-based anomaly detection and forensic analysis.

  Worked on other research problems which include profiling programs based on its DNS queries, protecting machine learning logics using hardware primitives (e.g., Intel SGX, ARM TrustZone).

- **IBM Korea** Mar 2001 - Aug. 2006.

  Advanced software engineer (Jul. 2004 - Aug. 2006): Advanced technical resource to handle critical projects or customer issues. Covered various IBM technologies; AIX operating system, IBM Java Virtual Machine (JVM), and high availability solutions.

Software specialist (Mar 2001 - Jun 2004): System engineer AIX operating systems and IBM's high availability solutions.

- **18 Medical Company, 8th U.S. Army** Jan. 1997 - Mar. 1999.

  Information management staff: IT infrastructure administrator 121 general hospital Yongsan U.S. army garrison, South Korea.

RESEARCH EXPERIENCE

- **Researcher, NEC Laboratories America**, Sep 2014 - Present.

  Protection of Machine Learning Algorithm with Hardware Assistance (Mar 2018 - Present): To protect the privacy of ML algorithms, I co-designed the ML platform that runs the algorithm inside Intel SGX and encrypts the channel for GPU communication.

  Application White-listing with Binary Triage (Mar 2018 - Present): I proposed and work on a white-listing solution that dynamically updates the reputation of application binary considering new metrics related to the trustworthiness of the binary.

  End-point Detection and Response for IoT devices (Sep 2017 - Present): In this project, we develop an EDR solution for IoT devices. To minimize the interference to the operation of IoT device, our approach only remains minimal workload to IoT devices and offload the rest to the cloud to protect IoT deployment with ML-based anomaly detection and support for forensic analysis.

  Enterprise security solution for ML detection and forensic support (Sep 2014 - Aug 2017): Autonomous Security intelligence (ASI) is an award-winning security solution that monitors and collects system activities from a large number of end-hosts to support ML-based anomaly detection and forensic analysis. I designed and developed auditing components for different end-hosts (Linux, OS X, IoT platforms). I was mainly in charge of the threat intelligence and the analysis the latest attack trend.

- **Research Assistant, Columbia University**, Sep 2008 - Aug 2014.

  IntFlow (Jun 2012 - Jul 2012): IntFlow is a compiler extension which uses static data-flow analysis to report integer errors, taking into account common developer practices to reduce false positives. IntFlow achieves 89% reduction in false positives compared to the Clang-based Integer Overflow Checker (IOC).

  ShadowReplica, (Jan 2012 - June 2013): ShadowReplica accelerates DFT and other shadow memory-based analyses, by decoupling analysis from execution and utilizing spare CPU cores to run them in parallel. DFT is run in parallel by a shadow thread that is spawned for each application thread, and the two communicate using a shared data structure. We avoid the problems suffered by previous approaches, by introducing an offline application analysis phase that utilizes both static and dynamic analysis approaches to generate optimized code for decoupling execution and implementing DFT, while it also minimizes the amount of information needed for communication between two threads.

  Taint Flow Algebra (TFA), (Jan 2010 - Dec 2011): To improve the efficiency of DFT, our approach separates the program logic from the corresponding tracking logic, extracting the semantics of the latter and abstracting them using an intermediate representation(IR) of Taint Flow Algebra. We then apply optimization techniques to eliminate redundant tracking logic and minimize interference with the target program.

  libdft, (Jan 2010 - Dec 2011): Dynamic data flow tracking (DFT) deals with tagging and tracking data of interest as they propagate during program execution. We presented libdft, a dynamic DFT framework that unlike previous work is at once fast, reusable, and works with commodity software and hardware. libdft provides an API for building DFT-enabled tools that work on unmodified binaries, running on common operating systems and hardware, thus facilitating research and rapid prototyping.

  Evaluation of control plane mechanisms in propagating security updates, (June 2008 - July 2010) The project measured and evaluated the reliability and performance trade-offs for a variety of control channel mechanisms that are suitable for coordinating large-scale collaborative defenses when under attack. Our results show that the performance and reliability characteristics change drastically when one evaluates the systems under attack by a sophisticated and targeted adversary.

TEACHING EXPERIENCE

- **Instructor.** Introduction to Programming (COMSW3101-003) Columbia University 2013 Fall
  14 students, course evaluation 4.45 / 5.0
- **Teaching Assistant** Artificial Intelligence (COMS W4701), Columbia University 2012 Spring
- **Teaching Assistant** Advanced Programming (COMS 3157), Columbia University 2010 Fall

PUBLICATIONS

PAPERS

1. Y. Tang, D. Li, Z. Li, M. Zhang, K. Jee, Z. Wu, J. Rhee, X. Xiao, F. Xu, Q. Li "NodeMerge: Template Based Efficient Data Reduction For Big-Data Causality Analysis" *Proc. of CCS*, November 2018
2. P. Gao, X. Xiao, D. Li, Z. Li, K. Jee, Z. Wu, C. Kim, S. R. Kulkarni, P. Mittal "SAQL: A Stream-based Query System for Real-Time Abnormal System Behavior Detection" *Proc. of USENIX Security*, August 2018
3. P. Gao, X. Xiao, Z. Li, K. Jee, F. Xu, S. R. Kulkarni, P. Mittal "AIQL: Enabling Efficient Attack Investigation from System Monitoring Data" *Proc. of USENIX ATC*, July 2018
4. Y. Liu, M. Zhang, D. Li, K. Jee, Z. Li, Z. Wu, J. Rhee, P. Mittal "Towards a Timely Causality Analysis for Enterprise Security" *Proc. of NDSS*, February 2018
5. Z. Xu, Z. Wu, Z. Li, K. Jee, J. Rhee, X. Xiao, F. Xu, H. Wang, G. Jiang "High fidelity data reduction for big data security dependency analyses" *Proc. of CCS*, October 2016.
6. M. Pomonis, T. Petsios, K. Jee, M. Polychronakis, A. D. Keromytis "IntFlow: improving the accuracy of arithmetic error detection using information flow tracking" *Proc. of ACSAC*, December 2014.
7. K. Jee, V. P. Kemerlis, A. D. Keromytis and G. Portokalidis, "ShadowReplica: Efficient Parallelization of Dynamic Data Flow Tracking" *Proc. of CCS*, October 2013.
8. V. P. Kemerlis, G. Portokalidis, K. Jee, and A. D. Keromytis, "libdft: Practical Dynamic Data Flow Tracking for Commodity Systems" *Proc. of VEE*, March 2012.
9. K. Jee, G. Portokalidis, V. P. Kemerlis, S. Ghosh, D. I. August, and A. D. Keromytis. "A General Approach for Efficiently Accelerating Software-based Dynamic Data Flow Tracking on Commodity Hardware" *Proc. of NDSS* February 2012.
10. K. Jee, S. Sidiroglou-Douskos, A. Stavrou, and A. D. Keromytis. "An Adversarial Evaluation of Network Signaling and Control Mechanisms" *Proc. of ICISC* December 2010.
11. "AIX Practical performance and tuning guide" published by IBM Press books, ISBN-0738491799 March 2005.

PAPERS UNDER REVIEW

12. Y. Sun, K. Jee, S. Sivakorn, Z. Li, C. Lumezanu, Z. Wu, L. Kort-Parn, J. Rhee, C. Kim, M. Chiang, P. Mittal "You are What You Query: Program DNS Behavior under Malware Injection" *In submission to IMC 2018*
13. S. Sivakorn , K. Jee, Y. Sun, L. Kort-Parn, Z. Li, Z. Wu, C. Lumezanu, L. Tang, D. Li "PDNS: Countering Malicious Processes with End-point DNS Monitoring" *In submission to NDSS 2019*
14. H. Zhang, L. Tang, J. Gao, Z. Chen, B. Zong, K. Jee and Z. Li "On Discovering Advanced Persistent Threats from Enterprise Monitoring Logs" *In submission to CIKM 2018*

PATENTS & APPLICATIONS

15. Automated blackbox inference of external origin user behavior Z Wu, J Rhee, J Yuseok, LI Zhichun, K Jee, G Jiang US Patent App. 15/652,796
16. Blackbox Program Privilege Flow Analysis with Inferred Program Behavior Context J Rhee, J Yuseok, LI Zhichun, K Jee, Z Wu, G Jiang US Patent App. 15/623,538
17. Fine-Grained Analysis and Prevention of Invalid Privilege Transitions J Rhee, J Yuseok, LI Zhichun, K Jee, Z Wu, G Jiang US Patent App. 15/623,589
18. Host level detect mechanism for malicious dns activities K Jee, LI Zhichun, G Jiang, L Korts-Parn, Z Wu, Y Sun, J Rhee US Patent App. 15/644,018
19. Extraction and comparison of hybrid program binary features J Rhee, LI Zhichun, Z Wu, K Jee, G Jiang US Patent App. 15/479,928
20. High Fidelity Data Reduction for System Dependency Analysis Z Wu, LI Zhichun, J Rhee, F Xu, G Jiang, K Jee, X Xiao, Z Xu US Patent App. 15/416,346

HONORS AND AWARDS

- CEATEC Award for successfully commercialization ASI
- Kaspersky student conference 2014 2nd place in North America round
- Awarded IBM Stock option (500 stocks)
- Selected for IBM top-talented group (resource pool for future executives) for being rated three PBC 1 (Personal Business Commitment) out of four personal performance evaluations
- Selected for employee education program with full tuition support granted for MS program in Columbia University
- The Army Commendation Medal for outstanding service, 1999 18th Medical Company, 8th US Army

TALKS

CONFERENCE TALKS

1. "ShadowReplica: Efficient Parallelization of Dynamic Data Flow Tracking" ACM CCS 2013, Berlin, Germany, Nov. 2013.
2. "A General Approach for Efficiently Accelerating Software-based Dynamic Data Flow Tracking on Commodity Hardware" NDSS 2012, San Diego, California, Feb. 2012.
3. "An Adversarial Evaluation of Network Signaling and Control Mechanisms" ICISC 2010, Seoul, South Korea, Dec. 2010.

INVITED TALKS

4. "ShadowReplica: Efficient Parallelization of Dynamic Data Flow Tracking", Security Group Seminar, Stevens Institute of Technology, Oct 2013.
5. "A General Approach for Efficiently Accelerating Software-based Dynamic Data Flow Tracking on Commodity Hardware", IBM Programming Language Day 2012, Jun. 2012, IBM Thomas J. Watson Research Center.
6. "A General Approach for Efficiently Accelerating Software-based Dynamic Data Flow Tracking on Commodity Hardware" Princeton University – Liberty Group, Mar 2011.

SERVICES

- Technical Program Committee Member, International Information Security Conference (ISC), Honolulu, HI, 2016
- External Reviewer (Conferences) - *CCS 2012, 2013, 2014, RAID 2014, 2015, CSET 20111, ESORICS 2011, ACNS 2011, CANS 2011, ICS 2012, IWSEC 2013, INFOCOM 2015, NDSS 2015*

Advising and Mentorship

MS Advising

- Daniel Song (MS student at Columbia University, currently Ph.d candidate at Columbia University)

  Project: Comparison study of Dynamic Binary Instrumentation (DBI) frameworks
- Mengqi Zhang (MS student Columbia University, currently software engineer at Facebook)

  Project: Compiler (LLVM) assisted program instrumentation and hardening
- Marios Pomonis, Theofilos Petsios (Ph.d candidate at Columbia University)

  Project: Arithmetic error detection using information flow tracking with compiler assisted program instrumentation.

Intern Advising

- Yasser Shalabi (Ph.d candidate at UIUC)

  Project: Fast and efficient system event collection from Linux kernel
- Yixin Sun (Ph.d candidate at Princeton University)

  Project: Analyzing Program DNS Behavior under Malware Injection

- Suphanee Sivakorn (Ph.d candidate at Columbia University)

  Project: System to Detect Malicious Processes with End-point DNS Monitoring
- Qi Wang (Ph.d candidate at UIUC)

  Project: End-point Machine Learning Detection for IoT Devices

LANGUAGES

Proficient in English, Korean.

REFERENCES

Dr. Angelos D. Keromytis
Program Manager, DARPA
Arlington, VA
phone: available on request
e-mail: available on request

Professor Salvatore S. Stolfo
Professor Columbia University
New York, NY
phone: available on request
e-mail: available on request

Professor Junfeng Yang
Associate Professor Columbia University
New York, NY
phone: available on request
e-mail: available on request

Dr. Geoff Jiang
Vice President, Technology
Ant Financial
Hangzhou, Zhejiang, China
phone: available on request
e-mail: available on request