Course Syllabus

Course	CS 7301.005.S20
Course Title	Recent Advances in Computing - Advanced Topics System Security
Professor	Kangkook Jee
Term	Spring 2020
Meetings	ECSS 3.910 Tues & Thurs 1:00 PM ~ 2:15 PM

Professor Contact Information

Office Phone	Primary Contact Phone Number
Office Location	ECSS 3.226
Email Address	kangkook.jee@utdallas.edu
Office Hours	TBD
Course Website	http://www.syssec.org/classes/cs7301.f20

Course Description

Systems and software running inside comprise underlying substrates in the recent advent of our computerized age. It is thus essential to keep these building blocks secure and reliable to build root for the chain of trust. In this course, we aim to learn traditional research problems in the system field and their solutions. Based on the historical review, we will study and understand various topics and problems in emerging system security areas.
This is a graduate-level course that mainly comprises three parts. The first part of the course will give a historical and principled overview of prominent attacks and their corresponding defensive measures. The course will review leading static and dynamic techniques that have been widely used in various defense approaches. The second part of the course will introduce emerging, frontier topics in system security research. We will cover various domains of provenance analysis, IoT, and ICS/CPS systems to learn how traditional approaches can be applied and further extended to new challenges. Lastly, the course will see how machine learning-based approaches can be applied to solve system security problems.
Throughout the semester, we will read and discuss seminal papers mainly from major security venues and other related fields. Based on gained knowledge in the class, students will work on class projects with an expectation of the mature deliverable and a conference-qualified report.

Student Learning Objectives/Outcomes

From the course, students will primarily learn the concepts, principles, and practices, widely used in traditional (or conventional) system security research. Students will also explore newly rising areas to learn how traditional wisdom can be used to address new security problems. Students will also learn how ML-based approaches can solve system security problems previously known-to-be difficult. After all, students will read 20+ papers and conduct in-class research projects to confirm their learning outcomes.

Required Textbooks and Materials

The materials for the courses include reading papers on various system security topics, provenance analysis, IoT, ICS/CPS systems, ML applications on security problems. Paper list will be posted on the course website.

Grading Policy

Paper summary (15 %)

We will study and discuss 2 ~ 3 papers per week. Students must read assigned papers before the class and submit the summary. Review submission will be due at midnight a day before the day. Most of the course readings will come from seminal papers in the field. Submission links to will are provided on the course page.

Paper presentation (20%)

Each student will be required to present 1-3 lectures of a paper assigned to the class, depending on the course enrollment. Students should prepare a detailed lecture with detailed slides. The slides will be distributed via eLearning. The course instructor will provide additional details on the presentation on the first day of class. All presenters must use the course template for either keynote or PowerPoint. Linux users can use the PowerPoint template with Open Office if they choose.

In-class discussion (15 %)

We can further enhance our understanding of a paper as we exchange opinions or understanding with peers. Many novel ideas were initiated and developed in the course of intense discussions. After the class presentation, we will try to dedicate 10 ~ 30 minutes for class discussion. Students are strongly encouraged to participate in class discussions actively.

Class project (50 %)

Building on knowledge gained in class, the primary deliverable from this course will be a course project. The students may work on research projects in groups and preferably complete a conference-quality report at the end of the semester. The paper's topic must be security-relevant, and the student(s) must be a lead author. Projects teams may include groups of up to two students; yet, the larger group size implies the greater project outcome. Students are encouraged to suggest and discuss their project ideas. I will soon upload the list of suggested project items. Details of the milestones and content will be given in class with the other project details. The instructor will advise each team/individual independently as needed. The project grade will be a combination of grades received for several milestone artifacts and the final conference-quality report. The project will be graded on novelty, depth, correctness, clarity of presentation, and effort. The projects will be presented in the final week.

Assignments & Academic Calendar

The follow schedule is subject to change.

Intro and overview

Week 1 (Jan 14 - Jan 16)	Logistics & Course Overview	Szekeres, L et al. **SoK - Eternal War in Memory.**
	System Security Landscape	Shoshitaishvili, Y. et al. **SOK - (State of) The Art of War - Offensive Techniques in Binary Analysis.**

System security fundamentals

Week 2 (Jan 21 - Jan 23)	Control Flow Defense	Abadi, M et al. **Control-flow integrity**
	Code Randomization	Koo, H. et al. **Compiler-Assisted Code Randomization.**
Week 3 (Jan 28 - Jan 3)	Data Flow Analysis	Jee, K. **A General Approach for Efficiently Accelerating Software-based Dynamic Data Flow Tracking on Commodity**
		Kwon, Y. et al. **LDX - Causality Inference by Lightweight Dual Execution.**

Provenance analysis

Week 4 (Feb 4 - Feb 6)	Provenance analysis	King, S. et al. **Backtracking intrusions**.
		Milajerdi, S. M. et al. **HOLMES - Real-Time APT Detection through Correlation of Suspicious Information Flows.**
Week 5 (Feb 11 - Feb 13)	Provenance-based anomaly detection	Yen, T.-F. et al. **Beehive - large-scale log analysis for detecting suspicious activity in enterprise networks.**
		Manzoor, E. et al. **Fast Memory-efficient Anomaly Detection in Streaming Heterogeneous Graphs**

IoT security

Week 6 (Feb 18 - Feb 20)	IoT security overview	Celik, Z. B. et al. **Program Analysis of Commodity IoT Applications for Security and Privacy - Challenges and Opportunities.**
	IoT device security	Cozzi, E. et al. **Understanding Linux Malware** Sun, Z. et.al. **OAT: Attesting Operation Integrity of Embedded Devices**

Week 7[1] (Feb 25 - Feb 27)	Smart Home Appliances	Alrawi, O. et al. **SoK - Security Evaluation of Home-Based IoT Deployments.**
		Wang, Q et al. **Fear and Logging in the Internet of Things**.
Week 8[2] (Mar 3 - Mar 5)	IoT Information Flow	Celik, Z. B. et al. **Sensitive Information Tracking in Commodity IoT**.
	ML approaches for IoT	Sikder, A. K. et al. **6thSense - A Context-aware Sensor-based Attack Detector for Smart Devices.** (Optional) Al-Garadi, M. et al. **A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security.**

ICS and CPS security

Week 9 (Mar 10 - Mar 12)	ICS security overview	McLaughlin, S. et al. **The Cybersecurity Landscape in Industrial Control Systems.** Douglas, B. et al. **The Fundamentals of Control Theory**
	Control theory and ICS attacks	Urbina, D. I. et al. **Limiting the Impact of Stealthy Attacks on Industrial Control Systems**
Week 10 (Mar 17 - Mar 19)	No class (Spring break)	
Week 11 (Mar 24 - Mar 26)	ICS code analysis	Keliris, A. et. al. **ICSREF: A Framework for Automated Reverse Engineering of Industrial Control Systems Binaries**
		Zhang, M. et al. **Towards Automated Safety Vetting of PLC Code in Real-World Plants.**
Week 12 (Mar 31 - Apr 2)	ICS malware and defenses	Garcia, L. et al. **Hey, My Malware Knows Physics! Attacking PLCs with Physical Model Aware Rootkit.**
		Chen, Y. et al. **Learning from Mutants - Using Code Mutation to Learn and Monitor Invariants of a Cyber-Physical System.**

ML for system security

Week 13 (Apr 7 - Jan 9)	Overview: Data mining and ML for cyber security	Buczak, A. L. et al. **A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection.**
Week 14 (Apr 14 - Apr 16)	Binary analysis with ML	Xu, X., Liu, C., Feng, Q., Yin, H., Song, L., & Song, D. (2017). **Neural Network-based Graph Embedding for Cross-Platform Binary Code Similarity Detection**
		Zhang, Z., Qi, P., & Wang, W. (2019). **Dynamic Malware Analysis with Feature Engineering and Feature Learning** Presented at the arXiv.org.

[1] NDSS week

[2] Mid-term week

Week 15 (Apr 21 - Apr 23)	ML-based IDS	Thomas, K., Grier, C., Ma, J., Paxson, V., & Song, D. (2011). **Design and Evaluation of a Real-Time URL Spam Filtering Service**
		Mirsky, Y. et al. **Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection**
Week 16 (Apr 28 - Apr 30)	More ML applications	Du, M. et al. **DeepLog - Anomaly Detection and Diagnosis from System Logs through Deep Learning.**
		She, D. et al. **Neutaint: Efficient Dynamic Taint Analysis with Neural Networks**

UT Dallas Syllabus Policies and Procedures

The information contained in the following link constitutes the University's policies and procedures segment of the course syllabus.

Please go to http://go.utdallas.edu/syllabus-policies for these policies.

The descriptions and timelines contained in this syllabus are subject to change at the discretion of the Professor.

Appendix

Intro and overview (1 week)

Module 1: Memory corruption and control hijacking attacks, buffer overflow, ROP, JIT-ROP (1 week)

Szekeres, L., Payer, M., Wei, T., & Song, D. (2013). **SoK - Eternal War in Memory.** (pp. 48–62). Presented at the IEEE Symposium on Security and Privacy, IEEE. http://doi.org/10.1109/SP.2013.13

Shoshitaishvili, Y., 0001, R. W., Salls, C., Stephens, N., Polino, M., Dutcher, A., et al. (2016). **SOK - (State of) The Art of War - Offensive Techniques in Binary Analysis.** *IEEE Symposium on Security and Privacy*.

System security fundamentals (3 weeks)

Module 1: Control Flow Integrity (½ week)

Zhang, M., & Sekar, R. (2013). **Control Flow Integrity for COTS Binaries.** Presented at the USENIX Security Symposium.

(Optional) Abadi, M., Budiu, M., Erlingsson, U. L., & Ligatti, J. (2005). **Control-flow integrity** (pp. 340–353). Presented at the ACM conference on Computer and communications security, New York, NY, USA: ACM. http://doi.org/10.1145/1102120.1102165

Module 2: Code Diversification / Randomization (½ week)

Koo, H., Chen, Y., Lu, L., Kemerlis, V. P., & Polychronakis, M. (2018)**. Compiler-Assisted Code Randomization.** (pp. 461–477). Presented at the IEEE Symposium on Security and Privacy, IEEE. http://doi.org/10.1109/SP.2018.00029

(Optional) Pappas, V., Polychronakis, M., & Keromytis, A. D. (2012). **Smashing the Gadgets - Hindering Return-Oriented Programming Using In-place Code Randomization.** Presented at the IEEE Symposium on Security and Privacy. http://doi.org/10.1109/SP.2012.41

Module 3: Information Flow Tracking (1 week)

(Optional) Kemerlis, V. P., Portokalidis, G., Jee, K., & Keromytis, A. D. (2012). **libdft: practical dynamic data flow tracking for commodity systems.** Presented at the ACM SIGPLAN/SIGOPS conference on Virtual Execution Environments, ACM Request Permissions. http://doi.org/10.1145/2151024.2151042

Jee, K., Portokalidis, G., Kemerlis, V. P., Ghosh, S., August, D. I., & Keromytis, A. D. (2012). **A General Approach for Efficiently Accelerating Software-based Dynamic Data Flow Tracking on Commodity Hardware.** *Network and Distributed System Security Symposium (NDSS)*.

Kwon, Y., Kim, D., Sumner, W. N., Kim, K., Saltaformaggio, B., 0001, X. Z., & Xu, D. (2016). **LDX - Causality Inference by Lightweight Dual Execution.** Presented at the International Conference on Architectural Support for Programming Languages and Operating Systems. http://doi.org/10.1145/2872362.2872395

Provenance analysis (3 weeks)

Module 1: Provenance analysis (1 week)

King, S. T., & Chen, P. M. (2003). **Backtracking intrusions**. Presented at the ACM Symposium on Operating Systems Principles. http://doi.org/10.1145/945445.945467

(Optional) Liu, Y., Zhang, M., Li, D., Jee, K., Li, Z., Wu, Z., et al. (2017). **Towards a Timely Causality Analysis for Enterprise Security** (pp. 1–15). Presented at the Network and Distributed System Security Symposium, Reston, VA: Internet Society. http://doi.org/10.14722/ndss.2018.23254

Milajerdi, S. M., Gjomemo, R., Eshete, B., Sekar, R., & Venkatakrishnan, V. N. (2019). **HOLMES - Real-Time APT Detection through Correlation of Suspicious Information Flows.** *IEEE Symposium on Security and Privacy*.

(Optional) Milajerdi, S. M., Eshete, B., Gjomemo, R., & Venkatakrishnan, V. N. (2019). **POIROT - Aligning Attack Behavior with Kernel Audit Records for Cyber Threat Hunting.** *ACM Conference on Computer and Communications Security*. http://doi.org/10.1145/3319535.3363217

Module 2: Provenance based Intrusion Detection (1 week)

Yen, T.-F., Oprea, A., Onarlioglu, K., Leetham, T., Robertson, W. K., Juels, A., & Kirda, E. (2013). **Beehive - large-scale log analysis for detecting suspicious activity in enterprise networks.** Presented at the Annual Computer Security Applications Conference. http://doi.org/10.1145/2523649.2523670

Manzoor, E., Milajerdi, S. M., & Akoglu, L. (2016). **Fast Memory-efficient Anomaly Detection in Streaming Heterogeneous Graphs** (pp. 1035–1044). Presented at the SIGKDD Conference on Knowledge Discovery and Data Mining, New York, New York, USA: ACM Press. http://doi.org/10.1145/2939672.2939783

IoT Security (3 weeks)

Module 1: IoT security introduction (½ week)

Celik, Z. B., Fernandes, E., Pauley, E., Tan, G., & McDaniel, P. (2019). **Program Analysis of Commodity IoT Applications for Security and Privacy - Challenges and Opportunities.** *ACM Comput. Surv.*, *52*(4), 1–30. http://doi.org/10.1145/3333501

Module 2: IoT device security (½ week)

Cozzi, E., Graziano, M., Fratantonio, Y., & Balzarotti, D. (2018). **Understanding Linux Malware** (pp. 161–175). Presented at the IEEE Symposium on Security and Privacy, IEEE. http://doi.org/10.1109/SP.2018.00054

Sun, Z., Feng, B., Lu, L., & Jha, S. (2020). **OAT: Attesting Operation Integrity of Embedded Devices** . Presented at the IEEE Symposium on Security and Privacy.

(Optional) Zheng, Y., Davanian, A., Yin, H., Song, C., Zhu, H., & Sun, L. (2019). **FIRM-AFL - High-Throughput Greybox Fuzzing of IoT Firmware via Augmented Process Emulation.** Presented at the USENIX Security Symposium.

Module 3: Smart Home Appliances (1 week)

Alrawi, O., Lever, C., Antonakakis, M., & Monrose, F. (2019). **SoK - Security Evaluation of Home-Based IoT Deployments.** *IEEE Symposium on Security and Privacy*, 1362–1380. http://doi.org/10.1109/SP.2019.00013

(Optional) Wang, Q., Hassan, W. U., Bates, A., & Gunter, C. (2017). **Fear and Logging in the Internet of Things**. Presented at the Network and Distributed System Security Symposium, Reston, VA: Internet Society. http://doi.org/10.14722/ndss.2018.23282

Module 4: Information flow at IoT (½ week)

Celik, Z. B., Babun, L., Sikder, A. K., Aksu, H., Tan, G., McDaniel, P., & Uluagac, A. S. (2018). **Sensitive Information Tracking in Commodity IoT**. *USENIX Security Symposium*.

(Optional) Fernandes, E., Paupore, J., Rahmati, A., Simionato, D., Conti, M., & Prakash, A. (2016). **FlowFence - Practical Data Protection for Emerging IoT Application Frameworks**. Presented at the USENIX Security Symposium.

Module 5: ML for IoT security (½ week)

Sikder, A. K., Aksu, H., & Uluagac, A. S. (2017). **6thSense - A Context-aware Sensor-based Attack Detector for Smart Devices.** Presented at the USENIX Security Symposium.

(Optional) Al-Garadi, M. A., Mohamed, A., Al-Ali, A., Du, X., & Guizani, M. (2018, July 29). **A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security.** *arXiv.org*.

(Optional) Module 6: Edge Computing Security

Konecný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). **Federated Learning - Strategies for Improving Communication Efficiency.** *CoRR*, *cs.LG*.

Yousefpour, A., Fung, C., Nguyen, T., Kadiyala, K., Jalali, F., Niakanlahiji, A., et al. (2019)**. All one needs to know about fog computing and related edge computing paradigms: A complete survey.** *Journal of Systems Architecture*, *98*, 289–330. http://doi.org/10.1016/j.sysarc.2019.02.009

ICS/CPS security (3 weeks)

Module 1: ICS security overview (½ week)

McLaughlin, S., Konstantinou, C., Wang, X., Davi, L., Sadeghi, A.-R., Maniatakos, M., & Karri, R. (2016). **The Cybersecurity Landscape in Industrial Control Systems.** *Proceedings of the IEEE*, *104*(5), 1039–1057. http://doi.org/10.1109/JPROC.2015.2512235

Module 2: Control theory and ICS attacks (½ week)

Douglas, B. (2019). **The Fundamentals of Control Theory** (pp. 1–160). Retrieved from https://drive.google.com/file/d/1LAjaDDViFG4H7dQ6PQVHo8XSQHS59GJf/view

Urbina, D. I., Giraldo, J. A., Cardenas, A. A., Tippenhauer, N. O., Valente, J., Faisal, M., et al. (2016). **Limiting the Impact of Stealthy Attacks on Industrial Control Systems** (pp. 1092–1105). Presented at the ACM conference on Computer and communications security, New York, New York, USA: ACM Press. http://doi.org/10.1145/2976749.2978388

Module 3: ICS code analysis (1 week)

Keliris, A., & Maniatakos, M. (2019). **ICSREF: A Framework for Automated Reverse Engineering of Industrial Control Systems Binaries** Presented at the Network and Distributed System Security Symposium, Reston, VA: Internet Society. http://doi.org/10.14722/ndss.2019.23271

Zhang, M., Chen, C.-Y., Kao, B.-C., Qamsane, Y., Shao, Y., Lin, Y., et al. (2019). **Towards Automated Safety Vetting of PLC Code in Real-World Plants.** (pp. 522–538). Presented at the IEEE Symposium on Security and Privacy, IEEE. http://doi.org/10.1109/SP.2019.00034

Moulde 4: ICS malware and defenses (1 week)

Garcia, L., Brasser, F., Cintuglu, M. H., Sadeghi, A.-R., Mohammed, O. A., & Zonouz, S. A. (2017). **Hey, My Malware Knows Physics! Attacking PLCs with Physical Model Aware Rootkit.** Presented at the Network and Distributed System Security Symposium. http://doi.org/10.14722/ndss.2017.23313

Chen, Y., Poskitt, C. M., & Sun, J. (2018). **Learning from Mutants - Using Code Mutation to Learn and Monitor Invariants of a Cyber-Physical System.** (pp. 648–660). Presented at the IEEE Symposium on Security and Privacy, IEEE. http://doi.org/10.1109/SP.2018.00016

ML for system security (4 weeks)

Module 1: Data mining and ML for cyber security IDS (1 week)

Buczak, A. L., & Guven, E. (2016). **A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection.** *IEEE Communications Surveys & Tutorials*, *18*(2), 1153–1176. http://doi.org/10.1109/COMST.2015.2494502

(Optional) Etalle, S. (2019). **Network Monitoring of Industrial Control Systems - The Lessons of SecurityMatters.** *Cps-Spc@Ccs*. http://doi.org/10.1145/3338499.3357354

Module 2: Binary analysis with ML (1 week)

Xu, X., Liu, C., Feng, Q., Yin, H., Song, L., & Song, D. (2017). **Neural Network-based Graph Embedding for Cross-Platform Binary Code Similarity Detection** (pp. 363–376). Presented at the the 2017 ACM SIGSAC Conference, New York, New York, USA: ACM Press. http://doi.org/10.1145/3133956.3134018

Zhang, Z., Qi, P., & Wang, W. (2019). **Dynamic Malware Analysis with Feature Engineering and Feature Learning** Presented at the arXiv.org.

(Optional) Rossow, C., Dietrich, C. J., Grier, C., Kreibich, C., Paxson, V., Pohlmann, N., et al. (2012). **Prudent Practices for Designing Malware Experiments: Status Quo and Outlook** (pp. 65–79). Presented at the IEEE Symposium on Security and Privacy, IEEE. http://doi.org/10.1109/SP.2012.14

Module 3: ML-based intrusion detection systems (1 week)

Thomas, K., Grier, C., Ma, J., Paxson, V., & Song, D. (2011). **Design and Evaluation of a Real-Time URL Spam Filtering Service** (pp. 447–462). Presented at the IEEE Symposium on Security and Privacy, IEEE. http://doi.org/10.1109/SP.2011.25

Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). **Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection** Presented at the Network and Distributed System Security Symposium, Reston, VA: Internet Society. http://doi.org/10.14722/ndss.2018.23204

Module 4: More ML applications to security (1 week)

Du, M., 0001, F. L., Zheng, G., & Srikumar, V. (2017). **DeepLog - Anomaly Detection and Diagnosis from System Logs through Deep Learning.** Presented at the ACM conference on Computer and communications security. http://doi.org/10.1145/3133956.3134015

She, D., Chen, Y., Shah, A., Ray, B., & Jana, S. (2019). **Neutaint: Efficient Dynamic Taint Analysis with Neural Networks** Presented at the IEEE Symposium on Security and Privacy.