Volume 6, Issue 1, Mar. 01, 2021

TC-CPS Newsletter

Technical Articles

Summary of Activities

Call for Contributions

IEEE TC-CPS Cyber-Physical Systems

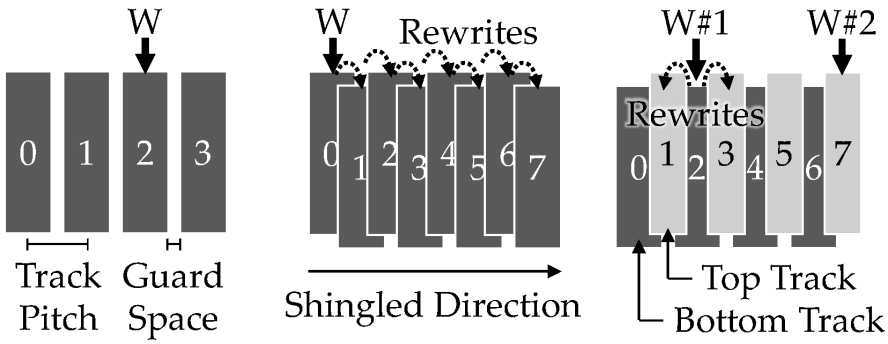
The Development of Hard Disk Drive Technology

Yuhong Liang, Ming-Chang Yang

Department of Computer Science and Engineering, The Chinese University of Hong Kong

1 Introduction

Hard disk drives (HDDs) have dominated the storage market for several decades because of its low cost-per-gigabyte and low bit-error-rate [1]. Although in recent years, solid-state drive (SSDs) have been penetrating into modern storage market, HDDs are still widely regarded as essential components for building cost-effective and reliable storage systems, such as data centers [2, 3, 4] and cloud storage [5], in the foreseeable future. Conventional Magnetic Recording (CMR) [6] is the most common type of track layout for HDDs. As illustrated in Figure 1(a), by introducing an empty space (i.e., *guard space* [7]) between any two consecutive tracks, CMR tracks can be written freely in any order, and each CMR track can be (re)written individually. However, the guard space may inevitably increase the distance (i.e., *track pitch*) between any two adjacent CMR tracks and limit the increasing of areal density capability (ADC) for HDDs. Recently, it has been reported that the CMR track layout has reached the bottleneck in providing higher areal density because of the superparamagnetic effect (SPE) [8]. To further enable the continued ADC growth for HDDs, various new magnetic recording technologies have been urgently investigated and proposed [9, 10, 11, 12]. Well known examples are *Singled Magnetic Recording (SMR)* [11, 7, 13], and *Interlaced Magnetic Recording (IMR)* [9].



(a) CMR (b) SMR (c) IMR

Figure 1: Track layouts of CMR, SMR, and IMR technologies.

2 Shingled Magnetic Recording (SMR) and Its Related Work

Shingled Magnetic Recording (SMR) [11, 7, 13] has drawn the most attention in the past few years, because of its strength in increasing the areal density with limited changes to the disk makeup [14]. As shown in Figure 1(b), the data are written into SMR tracks in a sequential fashion with tracks overlapped with the previous one(s), just like placing shingles on a roof. However, such a shingled design brings not only the narrower track pitch and higher areal density, but also the *sequential-write-constraint* [14]. That is, when updating data on any SMR track, the data stored

in the subsequent SMR track(s) will be overwritten and destroyed. Thus, to avoid losing data of subsequent tracks, "cascading" *track rewrites* (as known as *read-modify-write (RMW)* [15]) are necessary to re-write all the subsequent tracks containing valid data, even though such strategy may lead to severe "write amplification"[1] problem.

To mitigate the write amplification problem for SMR based HDDs, different models of SMR based HDDs are proposed, such as Drive-Managed SMR (DM-SMR) [13, 7, 17], Host-Managed SMR (HM-SMR) [18, 19], Host-Aware SMR (HA-SMR) [14, 20]. First, the Drive-Managed (DM) SMR drive model (DM-SMR) was presented as a traditional block device to cope with the sequential-write-constraint by the SMR drive itself so that the host system does not have to worry about the SMR characteristics [21]. Later, the Zoned Block Device (ZBC) standard[2] [22] is defined to 1) manage the SMR tracks in the unit of "zones" and 2) let the host system get involved in the storage management to better preserve the performance predictability of SMR drives. As presented by [13], SMR tracks can be organized into many "SMR zones" with the guard regions located in the middle of any two consecutive zones, so that the SMR sequential-write-constraint can be confined to a zone without crossing zone boundaries. According to the degree of strictness in complying with the sequential-write-constraint, the write pointer zone can be further classified into 1) *sequential-write-required zone* and 2) *sequential-write-preferred zone*. Based on the adoption of different zone types, SMR drives are further classified into two types. An SMR drive adopting sequential-write-required zones is referred to as the Host-Managed (HM) SMR drive model (HM-SMR), and the host system is strictly required to sequentially write data into any sequential-write-required zones of an HM-SMR drive, or an `ILLEGAL_REQUEST` will be reported by the drive. On the other hand, an SMR drive composed of sequential-write-preferred zones is regarded as the Host-Aware (HA) SMR drive model (HA-SMR), which usually incorporates a hidden Persistent Cache [13, 23] to alleviate the management burdens for the host system on handling the non-sequential writes.

To leverage the respective strengths of different SMR based HDD models, a good number of solutions are proposed at different software layers. For example, Singled Translation Layer (STL) [13] has been introduced at device-level. Virtual Persistent Cache [14] presents a driver-level solution. Besides, HiSMRfs [18], CaveatScriptor [19], and ShingledFS[24] are approaches at file-system-level, while SMRDB[25] is proposed at application level. Although all of these solutions effectively reduce the write amplification, SMR based HDDs still suffer serious write performance degradation, as compared with CMR based HDDs. As a result, in reality, not like CMR based HDDs could be widely deployed to a variety of computer systems, SMR based HDDs could be only applied to a limited number of applications or services, such as the archival or backup systems [26].

3 Interlaced Magnetic Recording (IMR) and Its Related Work

More recently, a new type of track layout, namely Interlaced Magnetic Recording (IMR) [9], has been proved for being capable of achieving 1) higher areal density and 2) much lower write amplification than the SMR track layout, as combined with the Heat-Assisted Magnetic Recording (HAMR) [27] and Multiple Sensor Magnetic Recording (MSMR) [28] technologies. To achieve a higher areal density for HDDs, IMR arranges tracks in an "interlaced" fashion shown in Figure 1(c). However, when updating data to the bottom tracks, valid data stored in the adjacent top tracks will be interfered and destroyed. Therefore, the *read-modify-write (RMW)* strategy has to be introduced to rewrite (at most two) adjacent top tracks. As the example shown in Figure 1(c), when the data stored in *track #2* are updated, the disk needs to 1) read the valid data from its adjacent top tracks (i.e., *track #1* and *track #3*), 2) perform data updates to *track #2*, and 3) re-write the two adjacent top tracks. Although the naïve RMW strategy can avoid data losses, it may inevitably introduce additional track rewrites (i.e., write amplification) and cause write performance degradation.

To date, only a few studies try to tackle the track rewriting issue of IMR. The first work is the *three-phase write management* method proposed by Gao *et al* [29, 30]. This approach allocates the disk space based on three phases of space usage. In the first phase, when the space usage is less than the total space of bottom tracks (roughly $0 \sim 50\%$ space usage), data will be only written into the bottom tracks without any track rewriting overhead. After all the

[1] The *write amplification* is defined as the ratio between the amount of data written to HDD and the amount of data requested by user [16].

[2] The Zoned Block Device is introduced by the T10 Technical Committee [22] to standardize the models and command sets to make the host system get involved in facilitating the zone-based storage media.

bottom tracks are consumed, the second phase starts to store data into every other top tracks (roughly $50 \sim 75\%$ space usage), while the remaining top tracks are only used to accommodate data in the third phase (roughly $75 \sim 100\%$ space usage). That is, during the second phase (resp. to third phase), updating a bottom track will require at most one (resp. to two) track rewrite(s).

Based on the three-phase write management method, Wu *et al.* further proposed a new data management design with two advanced techniques: namely *Top-Buffer* and *Block-Swap* [31]. Firstly, this new data management design manages the data in the unit/size of a logical block (e.g., 4 Kbytes or 8 Kbytes). That is, a block-level mapping table needs to be maintained to translate the logical block address (LBA) of data to the physical block address (PBA) in the IMR based HDDs. Secondly, the Top-Buffer technique allocates a few top tracks as write buffer to absorb the data updates and avoid track rewrites. Besides, when the buffer is full, the Top-Buffer technique proposes to write a whole track of data blocks from the buffer back to the corresponding track(s) at a time. Thirdly, the Block-Swap technique swaps the frequently-updated data in the bottom tracks with infrequently-updated data in the top tracks, when the space usage of the IMR based HDD is nearly full. Notably, for ease of our following discussion, the frequently-updated data (resp. to infrequently-updated data) will also be referred to as *hot data* (resp. to *cold data*).

More recently, Hajkazemi *et al.* [15] proposed several new mechanisms, such as *track flipping*, *selective track caching* and *dynamic track mapping* for IMR based HDDs. Different from managing data in block-level [31], these new write strategies adopt the memory-efficient track-based translation to manage the data in the unit/size of a track (e.g., 2 Mbytes). Specifically, the track flipping mechanism exchanges the hot data stored in bottom tracks with the data stored in its adjacent top tracks; the selective track caching mechanism redirects the "hottest" data of tracks to a reserved disk region (much like the persistent cache in an SMR based HDD [13]), which is composed of k bottom tracks; while the dynamic track mapping mechanism further exchanges the hot data of bottom tracks with the cold data of top tracks within a predefined range of tracks (called a band). Moreover, based on the description in [15], these mechanisms should be triggered in a periodical way.

4 Conclusion

The past decades have witnessed the tremendous success of Conventional Magnetic Recording (CMR) based Hard Disk Drives (HDDs) in data storage. To eliminate the bottleneck of CMR based HDDs in providing higher areal density, the emerging Shingled Magnetic Recording technology (SMR) and Interlaced Magnetic Recording technology (IMR) are capable of achieving higher areal density for HDDs with limited changes to disk makeup. Nevertheless, both SMR and IMR based HDDs may suffer write amplification problem brought from the special track layouts of SMR and IMR. To date, a number of studies have been proposed to mitigate the write amplification problem for both SMR based HDDs and IMR based HDDs.

References

[1] H. Jo, Y. Kwon, H. Kim, E. Seo, J. Lee, and S. Maeng, "Ssd-hdd-hybrid virtual disk in consolidated environments," in *European Conference on Parallel Processing*. Springer, 2009, pp. 375–384.

[2] Y. Li, X. Chen, N. Zheng, J. Hao, and T. Zhang, "An exploratory study on software-defined data center hard disk drives," *ACM Transactions on Storage (TOS)*, vol. 15, no. 3, pp. 1–22, 2019.

[3] S. Jaffer, S. Maneas, A. Hwang, and B. Schroeder, "The reliability of modern file systems in the face of ssd errors," *ACM Transactions on Storage (TOS)*, vol. 16, no. 1, pp. 1–28, 2020.

[4] E. Brewer, L. Ying, L. Greenfield, R. Cypher, and T. T'so, "Disks for data centers," 2016.

[5] Y. Yamato, "Cloud storage application area of hdd–ssd hybrid storage, distributed storage, and hdd storage," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 11, no. 5, pp. 674–675, 2016.

[6] K. Gao, "Drive architecture for hard disk drive," *IEEE Magnetics Letters*, vol. PP, pp. 1–1, 01 2018.

[7] W. He and D. H. Du, "Smart: An approach to shingled magnetic recording translation," in *15th {USENIX} Conference on File and Storage Technologies ({FAST} 17)*, 2017, pp. 121–134.

[8] A. Amer, J. A. Holliday, D. D. E. Long, E. L. Miller, J. F. Pâris, and T. Schwarz, "Data management and layout for shingled magnetic recording," *IEEE Transactions on Magnetics*, vol. 47, no. 10, pp. 3691–3697, 2011.

[9] E. Hwang, J. Park, R. Rauschmayer, and B. Wilson, "Interlaced magnetic recording," *IEEE Transactions on Magnetics*, vol. 53, no. 4, pp. 1–7, 2016.

[10] M. H. Kryder, E. C. Gage, T. W. McDaniel, W. A. Challener, R. E. Rottmayer, G. Ju, Y.-T. Hsia, and M. F. Erden, "Heat assisted magnetic recording," *Proceedings of the IEEE*, vol. 96, no. 11, pp. 1810–1835, 2008.

[11] S. Greaves, Y. Kanai, and H. Muraoka, "Shingled recording for 2–3 tbit/in^2," *IEEE Transactions on Magnetics*, vol. 45, no. 10, pp. 3823–3829, 2009.

[12] G. Mathew, E. Hwang, J. Park, G. Garfunkel, and D. Hu, "Capacity advantage of array-reader-based magnetic recording (armr) for next generation hard disk drives," *IEEE Transactions on Magnetics*, vol. 50, no. 3, pp. 155–161, 2014.

[13] A. Aghayev, M. Shafaei, and P. Desnoyers, "Skylight—a window on shingled disk operation," *ACM Transactions on Storage (TOS)*, vol. 11, no. 4, p. 16, 2015.

[14] M.-C. Yang, Y.-H. Chang, F. Wu, T.-W. Kuo, and D. H. Du, "On improving the write responsiveness for host-aware smr drives," *IEEE Transactions on Computers*, vol. 68, no. 1, pp. 111–124, 2018.

[15] M. H. Hajkazemi, A. N. Kulkarni, P. Desnoyers, and T. R. Feldman, "Track-based translation layers for interlaced magnetic recording," in *2019 {USENIX} Annual Technical Conference ({USENIX}{ATC} 19)*, 2019, pp. 821–832.

[16] L. Lu, T. S. Pillai, H. Gopalakrishnan, A. C. Arpaci-Dusseau, and R. H. Arpaci-Dusseau, "Wisckey: Separating keys from values in ssd-conscious storage," *ACM Transactions on Storage (TOS)*, vol. 13, no. 1, p. 5, 2017.

[17] M. Shafaei, M. H. Hajkazemi, P. Desnoyers, and A. Aghayev, "Modeling drive-managed smr performance," *ACM Transactions on Storage (TOS)*, vol. 13, no. 4, p. 38, 2017.

[18] C. Jin, W.-Y. Xi, Z.-Y. Ching, F. Huo, and C.-T. Lim, "Hismrfs: A high performance file system for shingled storage array," in *2014 30th Symposium on Mass Storage Systems and Technologies (MSST)*. IEEE, 2014, pp. 1–6.

[19] S. Kadekodi, S. Pimpale, and G. A. Gibson, "Caveat-scriptor: write anywhere shingled disks," in *7th {USENIX} Workshop on Hot Topics in Storage and File Systems (HotStorage 15)*, 2015.

[20] F. Wu, M.-C. Yang, Z. Fan, B. Zhang, X. Ge, and D. H. Du, "Evaluating host aware {SMR} drives," in *8th {USENIX} Workshop on Hot Topics in Storage and File Systems (HotStorage 16)*, 2016.

[21] Q. M. Le, K. SathyanarayanaRaju, A. Amer, and J. Holliday, "Workload impact on shingled write disks: All-writes can be alright," in *2011 IEEE 19th Annual International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems*. IEEE, 2011, pp. 444–446.

[22] "T10 technical committee of the international committee on infor- mation technology standards (incits)." *[Online]. Available: http://www.t10.org/.*

[23] F. Wu, Z. Fan, M.-C. Yang, B. Zhang, X. Ge, and D. H. Du, "Performance evaluation of host aware shingled magnetic recording (ha-smr) drives," *IEEE Transactions on Computers*, vol. 66, no. 11, pp. 1932–1945, 2017.

[24] A. Suresh, G. Gibson, and G. Ganger, "Shingled magnetic recording for big data applications," *Carnegie Mellon University Parallel Data Lab Technical Report CMU-PD L-12–105*, 2012.

[25] R. Pitchumani, J. Hughes, and E. L. Miller, "Smrdb: Key-value data store for shingled magnetic recording disks," 2015.

[26] R. Black, A. Donnelly, D. Harper, A. Ogus, and A. Rowstron, "Feeding the pelican: Using archival hard drives for cold storage racks," in *8th {USENIX} Workshop on Hot Topics in Storage and File Systems (HotStorage 16)*, 2016.

[27] S. Granz, W. Zhu, E. C. S. Seng, U. H. Kan, C. Rea, G. Ju, J.-U. Thiele, T. Rausch, and E. C. Gage, "Heat-assisted interlaced magnetic recording," *IEEE Transactions on Magnetics*, vol. 54, no. 2, pp. 1–4, 2017.

[28] S. Granz, J. Jury, C. Rea, G. Ju, J.-U. Thiele, T. Rausch, and E. C. Gage, "Areal density comparison between conventional, shingled, and interlaced heat-assisted magnetic recording with multiple sensor magnetic recording," *IEEE Transactions on Magnetics*, vol. 55, no. 3, pp. 1–3, 2018.

[29] K. Gao, W. Zhu, and E. Gage, "Write management for interlaced magnetic recording devices," Nov. 29 2016, US Patent 9,508,362.

[30] K. Gao, W. Zhu, and E. Gage, "Interlaced magnetic recording," Aug. 8 2017, US Patent 9,728,206.

[31] F. Wu, B. Zhang, Z. Cao, H. Wen, B. Li, J. Diehl, G. Wang, and D. H. Du, "Data management design for interlaced magnetic recording," in *10th {USENIX} Workshop on Hot Topics in Storage and File Systems (HotStorage 18)*, 2018.

An Overview of Physics-Informed Cyber Security Tools for Power Electronic Systems

Suman Rath[1] and Subham Sahoo[2]
[1]Veer Surendra Sai University of Technology Burla, India
[2]Aalborg University, Denmark

Abstract

Conversion-level power electronic devices are the key elements of a low-carbon-emitting world. This is mainly due to their role in integrating renewable sources of energy with the existing power transmission and distribution networks. However, reliability of such devices may be dented by the presence of an insecure cyber layer. An attacker can take advantage of the vulnerabilities in sensing and control layers to launch cyber attacks that can manipulate the working of large cyber-physical power electronic systems. This article presents an overview of such systems and gives an insight into how the cyber layer can be hijacked by attackers. Further, a detailed analysis of various physics-informed attack detection and mitigation strategies has been provided. Finally, a future course of action has been laid out towards the end of the article.

1 Introduction

Decarbonisation has been stressed upon as a key objective to be fulfilled in order to achieve the United Nations' Sustainable Development Goals (SDGs) [1]. As more countries strive to reduce their carbon footprint, renewable sources of energy come forth as a promising alternative to conventional fossil fuel-based power plants [2, 3]. Power electronics play a major role in aiding power conversion to facilitate energy exchange between renewables and the grid. Power electronics also enable power conversion for proper storage of energy which is essential as some renewables may not be able to generate sufficient power at all hours of the day [3]. Storage of excess energy allows its use during hours of peak demand. Thus, an increased share of renewables in the global energy market will also demand a proportionate increase in the share of power electronics in modern power systems.

Modern power electronics are cyber-physical systems which require continuous exchange of data for proper operation. Use of information and communication technologies (ICT) to enable better controllability of power electronics presents a duality as they also serve as portals using which intelligent hackers armed with the right tools and knowledge, can forcefully manipulate the system. This manipulation can be done in a multitude of ways including but not limited to, malware injection, data manipulation and denial of service. Consequences of such manipulations may range from mild to severe based on the target location and magnitude of attack. Extreme effects of such attacks can also result in prolonged power outages affecting millions of people and causing huge economic losses.

Possibility of such devastating consequences has motivated numerous attempts by researchers to develop mechanisms for detection and mitigation of cyber attacks in real-time. This article aims to contribute to the existing pool of research in three ways. First, it presents an overview of the architecture of cyber-physical power electronic systems focusing on ways in which the cyber layer can be hacked. Second, it analyses some latest physics-informed attack detection as well as mitigation techniques. A comparison of such strategies is important in order to gauge their efficacy on the basis of a common, standardized set of criteria. Third, this article also charts a possible plan of action for future researchers who choose to work in this domain.

2 General Architecture and Cyber Security

Power electronics are responsible for a number of functions including voltage regulation, frequency control and regulation of load sharing in a multi-source power system. Need for better monitoring, quick diagnosis and robust controllability has transformed power electronic systems into cyber-physical systems which require efficient data exchange networks for effective operation. These data exchange networks are responsible for communicating sensor inputs to the controllers that in turn, issue commands to regulate working of the physical devices. These networks also facilitate communication to enable synchronization among multiple converters. The architecture of a typical power-electronics dependent power system is shown in Figure 2. The system can be compared to the human body with the cyber layer acting as the brain that receives sensory inputs and controls the physical layer. This analogy

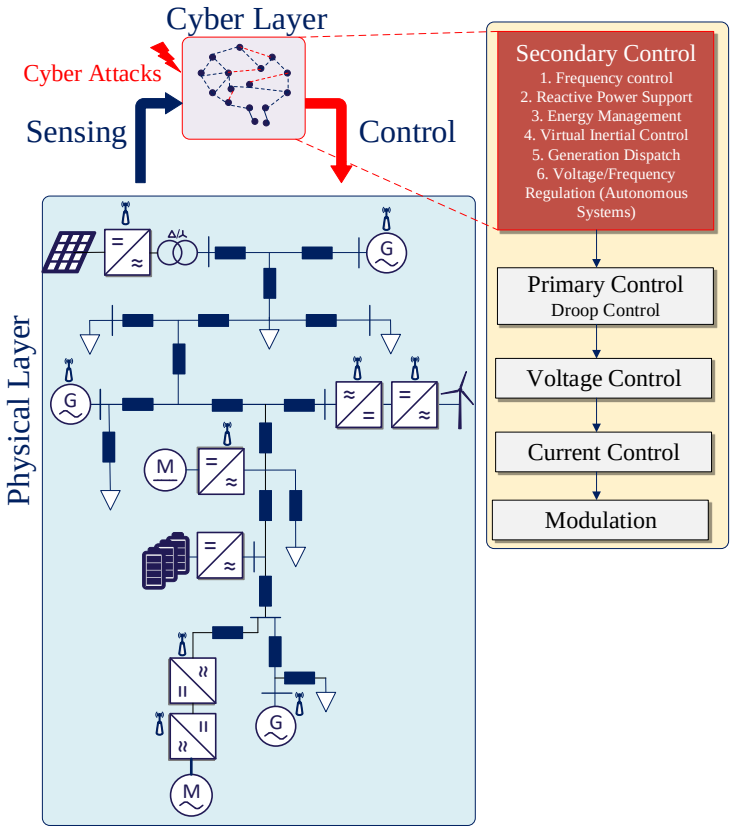


Figure 1: Architecture of a typical cyber-physical power electronics-dependent power system.

marks the position of the cyber layer as the most vital organ in a cyber-physical system. A breached cyber layer can be exploited to create voltage and frequency instability in the system which may even lead to blackouts. Manipulations targeting power converters may also be aimed towards affecting load sharing among multiple sources in a power system to disrupt optimal scheduling.

As shown in Figure 2, an attacker attempts to manipulate the cyber layer in three ways: (1) Sensor Manipulation, (2) Controller Manipulation and (3) Communication Network Manipulation. Sensor manipulation usually requires physical access to the sensors to enable tampering. However, Internet-of-Things (IoT)-enabled sensors can also be remotely hijacked. A malfunctioning sensor may lead to incorrect measurements that are further processed by controllers to generate erroneous commands. Controller manipulation is usually done through injection of malware into the host computer. Malware like Trojan can enable unauthorized remote accessibility of the host computer and hence, the associated controller. Communication network manipulation generally refers to manipulation of data flowing through the network. Manipulation of data stream through false data injection, replay attack and denial-of-service attack will lead to fraudulent inputs that will trick the controller into producing erroneous command signals. There are basically two ways to manipulate the communication network. The first technique involves injection of a remote controlled alien device into the network hardware consisting of transmitters, receivers and repeaters. The second technique is only applicable to wireless communication networks. It involves the use of sophisticated software like network analyzer tools to monitor wireless traffic and alter its contents. The next section presents a discussion on some of the latest cyber-breach detection and attack mitigation techniques to gauge their effectiveness on the basis of a standard evaluation scale.

3 Physics-Informed Attack Detection and Mitigation

Over the last decade, the research community has proposed numerous attack detection and mitigation strategies to secure cyber-physical systems. As shown in Figure 3, all such strategies can be broadly classified into two types:

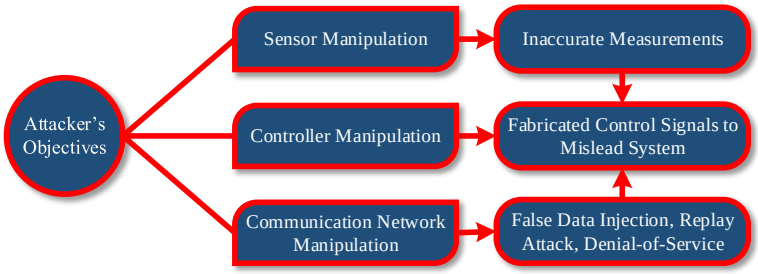


Figure 2: Objectives that an attacker tries to accomplish.

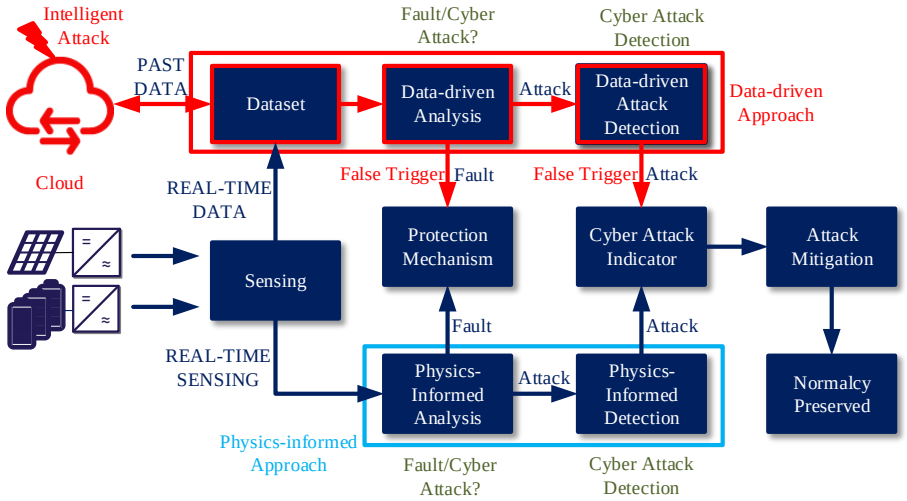


Figure 3: Data-driven and physics-informed approaches to cyber attack detection and mitigation.

data-driven and physics-informed approaches. Data-driven techniques extract data from the system and scrutinise it to detect abnormalities. This scrutiny generally involves examination of present datasets on the basis of past reference datasets. These techniques are often unreliable due to the possibility of data-based errors. Moreover, an intelligent hacker will always try to manipulate the past reference data to mask the attack. Hence, data-driven detection strategies are not preferable in a cyber-physical system. On the other hand, physics-informed techniques use their sensitivity to physical violations for detection of cyber attacks on the system. These methods do not depend on data verification and hence, are immune to data-based errors. This section presents a discussion on such physics-informed strategies by analysing them on the basis of features and functions of an ideal detection-cum-mitigation strategy as shown in Figure 4.

A strategy to detect malicious intrusions aiming to disturb the load sharing among sources in an AC microgrid has been mentioned in [4]. The strategy also classifies the manipulation attempts into fault attacks and random attacks. The proposed mitigation strategy eliminates any possible error in the optimal economic dispatch due to fault attacks by generating an update control signal and removes effects of random attacks by using local measurements from the pre-attack time period to estimate the incremental cost ensuring continuation of optimal operation.

Zhou *et al.* [5] has used a switching frequency-based mechanism to detect cyber attacks on communication links and controllers of a multi-source islanded microgrid. After detection, a mitigation strategy isolates the breached unit from the communication network topology to eliminate chances of attack propagation. In [6], cyber attacks on the communication network feeding inputs to the local controllers of Distributed Generators (DGs) in an autonomous

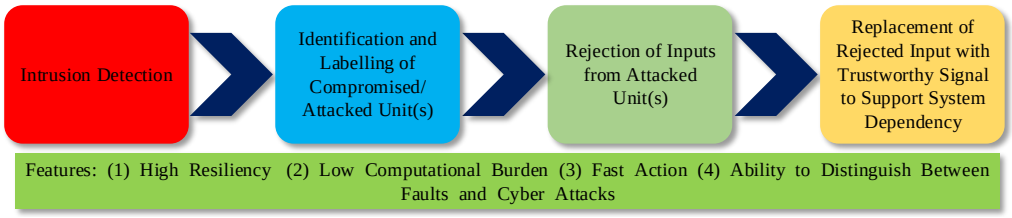


Figure 4: Functions and features of an ideal attack detection-cum-mitigation strategy.

AC microgrid have been detected through sensing of alteration in encryption and time stamps. After detection, a Link Error Counter assigned to each communication link is used to identify the attacked unit(s). A positive counter value triggers the mitigation process which reconfigures the secondary controller so as to authorize an alternate network topology and eliminate system dependency on the attacked frequency, real power or reactive power signal(s). This mechanism has a case-dependent resilience capability based on the number of DGs and links in the network and it exerts a medium computational burden on the system. A limitation of the proposed methods in [5] and [6] is their inability to detect or mitigate sensor manipulations.

Sahoo *et al.* [7] presents an event-driven mechanism to detect and mitigate cyber attacks in distributed AC microgrids. The proposed mechanism uses an asynchrony index-based metric to detect the presence of alien inputs in the frequency control signal. After detection, an authentication label is generated for the attacked agent which triggers the mitigation mechanism. The attacked signal is then reconstructed to support system dependency. A noteworthy feature of this technique is its high resiliency (N-1). This technique exerts a significantly low computational burden on the system due to the absence of complex mathematical operations. A similar event-driven strategy with high resiliency and low computational burden has been proposed in [8] to detect and mitigate attacks in DC microgrids. The proposed strategy is sensitive to stealth attacks designed to alter voltage and current signals. It uses a detection threshold to identify an attacked signal and mitigates any possible effects of manipulation by reconstructing trustworthy inputs to replace the attacked signals. A technique to detect and mitigate man-in-the-middle attacks targeting the communication links of a distributed DC microgrid has been presented in [9]. The detection mechanism uses a set of metrics to detect foreign elements in measurement data. Post-detection, the attacked unit is identified and a multilayer paradigm is used to extract a trustworthy version of the compromised measurement signal in order to mitigate possible effects of the cyber attack.

4 Conclusion and Future Work

This article provides an overview of cyber-physical power electronic systems especially focusing on elements which an attacker tries to hijack in order to manipulate the system i.e., sensors, communication networks and controllers. Significance of the role played by these elements to preserve system stability has been mentioned. A comparison has been drawn between data-driven and physics-informed approaches for cyber attack detection and mitigation to demonstrate the latter's superiority over the former. An analysis of some detection and mitigation strategies has been discussed on the basis of desirable features in an ideal strategy. Future work in this domain should aim to design detection-cum-mitigation strategies with high resilience capability and which exert only minor computational burden on the system. Future research should also strive to reduce the attack detection time and ensure quick mitigation.

References

[1] J. D. Sachs, G. Schmidt-Traub, M. Mazzucato, D. Messner, N. Nakicenovic and J. Rockström. Six transformations to achieve the sustainable development goals. *Nature Sustainability*, 2, pp. 805-814, Sep. 2019.

[2] D. Ciupăgeanu, G. Lăzăroiu and M. Tîrşu. Carbon dioxide emissions reduction by renewable energy employment in Romania. In *Proc. of International Conference on Electromechanical and Power Systems*, 2017, pp. 281-285.

[3] H. Gokul *et al*. Energy management and economical analysis of solar energy system for industrial applications. In *Proc. of International Conference on TAP Energy*, Kollam, India, 2017, pp. 1-6.

[4] S. Sahoo and J. C. -H. Peng. A localized event-driven resilient mechanism for cooperative microgrid against data integrity attacks. *IEEE Transactions on Cybernetics*, to be published, doi: 10.1109/TCYB.2020.2989225.

[5] Q. Zhou, M. Shahidehpour, A. Alabdulwahab and A. Abusorrah. A cyber-attack resilient distributed control strategy in islanded microgrids. *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 3690-3701, Sept. 2020.

[6] S. Rath, D. Pal, P. S. Sharma and B. K. Panigrahi. A cyber-secure distributed control architecture for autonomous AC microgrid. *IEEE Systems Journal*, to be published, doi: 10.1109/JSYST.2020.3020968.

[7] S. Sahoo, Y. Yang and F. Blaabjerg. Resilient synchronization strategy for AC microgrids under cyber attacks. *IEEE Transactions on Power Electronics*, vol. 36, no. 1, pp. 73-77, Jan. 2021.

[8] S. Sahoo, T. Dragičević and F. Blaabjerg. An Event-Driven Resilient Control Strategy for DC Microgrids. *IEEE Transactions on Power Electronics*, vol. 35, no. 12, pp. 13714-13724, Dec. 2020.

[9] S. Sahoo, T. Dragičević and F. Blaabjerg. Multilayer Resilience Paradigm Against Cyber Attacks in DC Microgrids. *IEEE Transactions on Power Electronics*, vol. 36, no. 3, pp. 2522-2532, March 2021.

A Defense Framework Against Long-Wire-Based Secret Leakage in Cloud-FPGA

Shijin Duan, Yukui Luo, Xiaolin Xu
Northeastern University

Abstract

The development of Cyber-Physical Systems (CPS) has been significantly empowered by cloud computing, i.e., gaining more computing capabilities from the cloud hardware infrastructures to process and visualize big data. This process also greatly accelerates compute-intensive and distributed applications. With the emerging cloud-computing development, FPGAs are being integrated with cloud servers for higher performance. Within cloud-FPGAs, multiple users can share the hardware FPGA resources to execute their own applications, which unfortunately incurs security concerns. It has been demonstrated that the capacitive crosstalk between FPGA long-wires can be a side-channel to extract secret information, giving adversaries the opportunity to implement crosstalk-based side-channel attacks. This work proposes a defense framework leveraging proper placement, routing, and obfuscation on FPGA resources to mitigate the secret leakage on long-wires. As a user-friendly defense strategy, this framework focuses on protecting the security-sensitive instances meanwhile possessing good compatibility with current development tools, such as ISE and Vivado.

1 Introduction

Due to the high computing-capability requirements in modern applications, accelerators such as Field Programmable Gate Arrays (FPGAs) are raised for large-scale parallel computation. As the in-real-time reaction demand is emerging and the application of artificial intelligence is making the high computing-capability urgent, FPGAs have been adopted for specifically massive computing or integrated into various platforms for performance acceleration [1] [2]. This implementation is quite suitable to Cyber-Physical System (CPS) with consideration of multiple aspects including performance, security, flexibility, and etc [3]. Besides, enabling FPGA on cloud [4] and realizing FPGA virtualization [5] are researched in recent years as well. Motivated by the market growth of cloud service, significant research and engineering efforts from various communities have been invested in advancing the performance of cloud-FPGAs, such as virtualizing the FPGA hardware resources for multiple users in parallel [5]. Combining with the advantages of the cloud-computing, several works have been explored on the Cloud-based CPS [6] [7]. Although the utilization of cloud-FPGA can significantly improve the performance and hardware utilization efficiency, security issue is also a critical consideration in CPS. For example, some recent works have demonstrated that the capacitive crosstalk between the long-wires on an FPGA chip can be used as a new side-channel to extract secret information [8, 9]. We assume two close parallel long-wires are *receiver* and *transmitter*, between which the equivalent capacitance cannot be neglected because of the short distance between wires in nowadays advanced FPGAs. It is shown that if the logic state on the *transmitter* is logic `0` and the *receiver* is transmitting a rising edge, it will take longer to make rising edge reach to the end of the *receiver*, compared to the case that the *transmitter* is carrying logic `1`[10]. Then the attacker can utilize this deviation to infer the state of the *transmitter* by measuring the propagation delay on the *receiver*. This phenomenon has been validated with various commodity FPGAs including Virtex-5, Virtex-6, and Artix-7 from Xilinx, as well as Cyclone-IV E, Cyclone-IV GX, and Stratix-V GX from Intel.

In response to this vulnerability, we propose a defense framework to mitigate the secret leakage on the long-wires in cloud-FPGAs. This framework optimizes the placement, routing, and leveraging obfuscation of security-sensitive FPGA components to eliminate the side-channel leakage caused by the crosstalk. It can prioritize placing and routing high-security components in the central area of design against other components and users, ensuring the security-sensitive wires will not be routed adjacently with possible malicious wires within other modules. Also, the long-wires are avoided to be used, so that the capacitive crosstalk can also be reduced between wires.

2 Defense Framework

We present the proposed defense framework in this section. The key idea of this framework is to eliminate the usage of long-wires and make the security-sensitive components wrapped in the central of the design implementations, so that the adjacent wires of these components are occupied by other insensitive resources in the same module while not being routed by external malicious modules. The schematic of the defense framework is shown in Fig. 1. The netlist

of targeted design and the sensitive net keywords are fed in the isolation algorithm, which will derive a constraint file with the placing arrangement. The *automatic isolation* will extract the components in the netlist as `resource list` and `connection list`, respectively. The sensitive keywords, representing the sensitive nets, are given by the designer as well for the *automatic isolation* to generate the hierarchical Physical Blocks (Pblocks) as the placing arrangement. Added with the long-wire obfuscation protecting those long-wires that cannot be eliminated, the constraint file is loaded in the design tools to guide the implementation against the long-wire information leakage. One great advantage of this defense framework is that it only requires the netlist of the design module, targeting the post-synthesis process. Hence, it keeps the majority privacy of the design by not requiring the configuration details of every Look-Up Tables (LUTs).

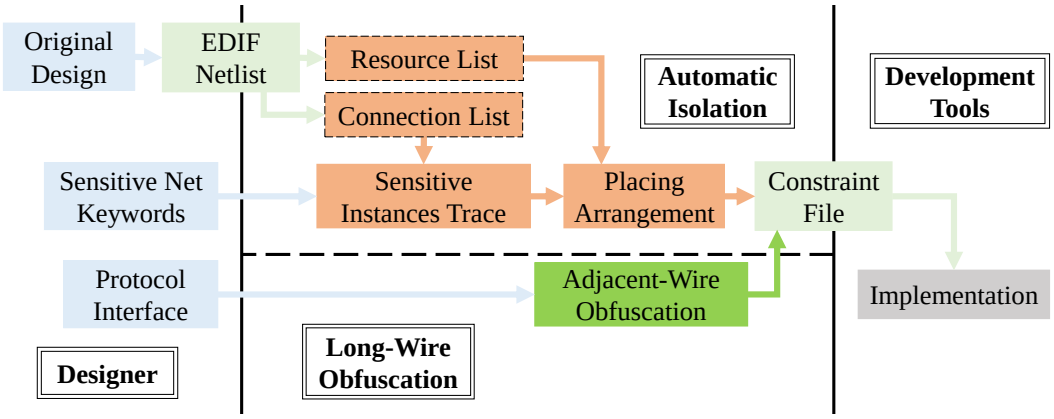


Figure 1: The defense framework schematic.

2.1 Automatic Isolation Algorithm

The *automatic isolation* algorithm can ensure that the long-wires are not utilized within the security-sensitive components, by placing them closely in the central area of the entire module. Further, central security-sensitive components will be wrapped with other insensitive components, in order to prevent other parts and tenants from placing instances and wires near them. To implement this algorithm, users only need to provide the netlist and the sensitive keyword for instance search.

The netlist is given as the Electronic Design Interchange Format (EDIF) file. Then, two primitive lists (`resource list` and `connection list`) are generated from the netlist. The `resource list` includes standard instances like LUT, DFF, and shift-register, which are used to estimate the usage of hardware resources. The `connection list` includes all net names and the connection between instances. Assuming the *automatic isolation* algorithm identifies m security-sensitive instances with n security keywords, we create a $m \times m$ weight matrix (**WM**), where the numbers stored denote the overall connections between any two critical instances, to evaluate the security-levels of different instances. Specifically, if the i^{th} row and j^{th} column of **WM** is 1, we can derive that there is a security-sensitive net connected between these two instances; the weight values for insensitive or unconnected instances are assigned as 0. The **WM** generation algorithm is shown in Alg. 1. The weight value for the j^{th} instance will be formulated by scanning all the n sensitive net keywords and summing the numbers in the row of **WM**. At the end of the *automatic isolation* algorithm, we return the Weight of the m critical instances: $\mathbf{W} = \{w_1, w_2, ..., w_m\}$, where w_j is the weight value for the j^{th} instance. We state that one instance is more security-sensitive if it is connected with more sensitive-nets, i.e. having a larger weight value.

The security-sensitive instances are sorted based on their weight values from high to low, and this ranking information is used to build a hierarchy for the placement and routing of the sensitive instances. We divide the hierarchy into s levels ($s = 5$ in our case) and put the most security-sensitive components (i.e. ones with highest weight values) in the central level, while the most insensitive components are placed in the outermost levels. Then, we spirally place the s levels in Pblocks, where the most central level will be in the center of the arrangement, wrapped by the outer less critical levels. When we do the spiral placement and routing, we first estimate the overhead

Algorithm 1 Calculate the weight for m critical instances: $\mathbf{W} = \{w_1, w_2, ...w_m\}$

Require: An all-zero weight matrix (**WM**) of dimension $m \times m$
Require: n critical net names $\{net_1, net_2, ..net_n\}$
Ensure: Weight of the m critical instances: $\mathbf{W} = \{w_1, w_2, ...w_m\}$

for (i$\in 1,2,3,...n$) **do**
  $P_i \leftarrow$ connection-ship between m instances via net_i
  for (j$\in 1,2,3,...m$) **do**
    for (k$\in 1,2,3,...m$) **do**
      if $\mathrm{wm}_{jk} \in P_i$ **then**
        $\mathrm{wm}_{jk} = \mathrm{wm}_{jk} + 1$
      end if
    end for
  end for
end for
for (j$\in 1,2,3,...m$) **do**
  $w_j = \sum_{l=1}^{m} \mathrm{wm}_{jl}$
end for
Return **W**

of each level. We assume the instances are placed in squared Pblocks, so that the starting point and endpoint of each security level can be easily determined for the Pblock constraints.

2.2 Long-Wire Obfuscation

Besides the general sensitive instances, like LUT and DFF, various data or communication interfaces and modules are commonly used in the FPGA implementations. Since these communication instances usually connect the module to external I/O ports on FPGA with a long-distance, they must utilize long-wires to transmit data, making secret information leakage possible as well. These long-wires cannot be eliminated because the peripheral I/O ports are fixed, which cannot be wrapped in the central area of Pblocks. To mitigate this vulnerability, we propose another method: *long-wire obfuscation* to reduce the crosstalk-based side-channel leakage on such interfaces. For an interface instantiated with long-wires, the *long-wire obfuscation* method will add obfuscation to adjacent long-wires of the communication interface. Specifically, the two long-wires on its left and right sides are occupied with obfuscation long-wires, transmitting constant logic(`0`/`1`) or random noise.

3 Experimental Validation

We validate the proposed defense framework using the open-source AES design [11] on a Xilinx Artix-7 FPGA, while implementing the design on both ISE 14.7 and Vivado 2018.3. For analyzing the defense framework more specifically, we evaluate the *automatic isolation* algorithm and *long-wire obfuscation* individually.

We first validate the *automatic isolation* algorithm on this AES design. In our validation, we mark the inputs of S-box (with net keyword "sr_out") as sensitive and applied the *automatic isolation* algorithm. The isolation algorithm mark 34 instances as security-sensitive that can affect the S-Box inputs. Following the flow in Fig. 1, these security-sensitive instances (marked with red) are placed and routed in the central region of the entire AES implementation, as shown in Fig.2. Since other hardware components in the AES design surrounds the sensitive instances, therefore, it is infeasible for other tenants to approach and attack them via crosstalk.

Then, the *long-wire obfuscation* method is evaluated based on the AES design. To explicitly show the obfuscation method, we fine-tune the I/O structure by building UART to transfer data rather than storing them in Block RAMs at first. We use UART as the interface of the AES implementation, through which the secret information (i.e., key or plaintext) is transmitted. We occupy the adjacent long-wires (i.e., the left-side and right-side long-wires) of the UART_RX interface, transmitting the constant logic `1`, logic `0`, or random signals through these two adjacent long-wires. The experimental results for using different obfuscation methods are shown in Tab. 1. Without any protection,

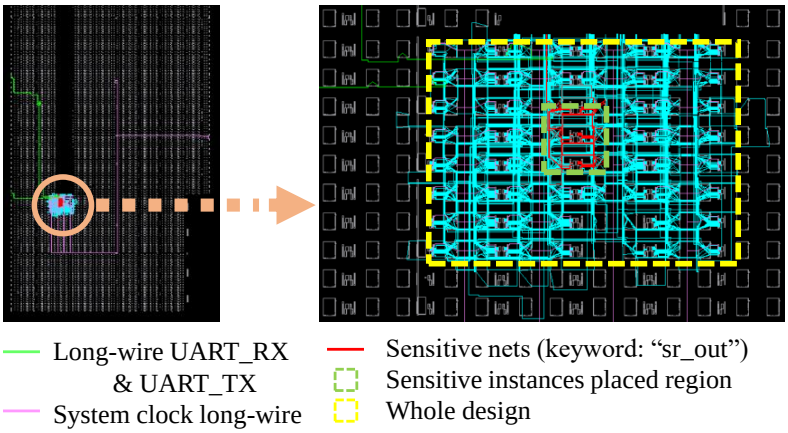


Figure 2: Experimental results for *automatic isolation* algorithm. The security-sensitive instances connected with sensitive net keyword “sr_out” are placed in the central region, while the rest insensitive instance are placed in the outer area.

Protection method		Prediction rate	
		Secret	Obfuscation logic
No protection		81.7%	X
Long-wire obfuscation	0	50.2%	50.3%
	1	50.45%	51.05%
	Random	50.45%	83.2%

Table 1: The experimental results for different long-wire obfuscation methods.

81.7% of the samples can be correctly predicted. When the security-sensitive long-wire, UART_RX, is protected by obfuscation long-wires, the prediction rate of secret information is reduced to $\sim$ 50%, which is quite close to the random guess. Interestingly, the side-channel leakage from crosstalk achieves a good prediction rate (83.2%) on the random obfuscation-wires. However, this randomly generated signal does not leak any secret information to the tenants.

4 Conclusion

While more implementations, such as Cyber-Physical System, are applied in cloud accelerators, the security issue in the cloud server is getting attention. In this work, we target a long-wire-based vulnerability found on cloud-FPGA and present a hardware defense framework against the side-channel information leakage leveraging long-wire capacitive crosstalk. The defense framework is composed of two algorithms: (1) the *automatic isolation* algorithm that can automatically identify and place the most security-sensitive components in the central area of the targeted design to protect them from eavesdropping by adjacent long-wires, and (2) the *long-wire obfuscation* method that can protect the communication interfaces with long-wires not-able-to be eliminated by isolation algorithm. We implemented and validated the performance of this defense framework with an Artix-7 FPGA. The experimental results validate the feasibility of these two methods, which significantly reduce the information leakage while protecting the sensitive nets. As a developer-friendly defense framework, it only needs the netlist of the targeted design so that cooperating with various FPGA development tools like ISE and Vivado.

References

[1] István, Zsolt et al. Providing multi-tenant services with FPGAs: Case study on a key-value store. In *Proc. of FPL*, pages 119–1195, 2018.

[2] Elnaggar, Rana et al. Multi-Tenant FPGA-based Reconfigurable Systems: Attacks and Defenses. In *Proc. of DATE*, pages 7–12, 2019.

[3] Rodr'ıguez, Alfonso et al. Fpga-based high-performance embedded systems for adaptive edge computing in cyber-physical systems: The artico3 framework. In *Sensors*, pages 1877, 2018.

[4] Chen, Fei et al. Enabling FPGAs in the cloud. In *Proc. of the 11th ACM Conference on Computing Frontiers*, pages 1–10, 2014.

[5] Yue Zha et al. Virtualizing FPGAs in the Cloud. In *Proc. of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, 2020.

[6] Villalonga, Alberto et al. Cloud-based industrial cyber–physical system for data-driven reasoning: A review and use case on an industry 4.0 pilot line In IEEE Transactions on Industrial Informatics, pages 5975–5984, 2020.

[7] Rajrup Ghosh et al. Distributed Scheduling of Event Analytics across Edge and Cloud. In ACM Trans. Cyber-Phys. Syst. 2, 4, Article 24 (September 2018), 28 pages, 2018.

[8] Giechaskiel, Ilias et al. Leaky wires: Information leakage and covert communication between FPGA long wires. In *Proc. of the 2018 on Asia Conference on Computer and Communications Security*, pages 15–27, 2018.

[9] Ramesh, Chethan et al. FPGA side channel attacks without physical access. In *IEEE 26th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pages 45–52, 2018.

[10] Luo, Yukui et al. Hill: A hardware isolation framework against information leakage on multi-tenant fpga long-wires. In International Conference on Field-Programmable Technology (ICFPT) pp. 331-334, 2019.

[11] 8bit datapath hardware implementation of AES. https://github.com/ChengluJin/8bit_datapath_AES.

Aerial Manipulation Using a Novel Unmanned Aerial Vehicle Cyber-Physical System

Caiwu Ding[1], Hongwu Peng[2], Lu Lu[1], and Caiwen Ding[2]
[1]Department of Mechanical and Industrial Engineering, New Jersey Institute of Technology
[2]Department of Computer Science and Engineering, University of Connecticut

Abstract

Unmanned Aerial Vehicles(UAVs) are attaining more and more maneuverability and sensory ability as a promising teleoperation platform for intelligent interaction with the environments. This work presents a novel 5-degree-of-freedom (DoF) unmanned aerial vehicle (UAV) cyber-physical system for aerial manipulation. This UAV's body is capable of exerting powerful propulsion force in the longitudinal direction, decoupling the translational dynamics and the rotational dynamics on the longitudinal plane. A high-level impedance control law is proposed to drive the vehicle for trajectory tracking and interaction with the environments. In addition, a vision-based real-time target identification and tracking method integrating a YOLO v3 real-time object detector with feature tracking, and morphological operations is proposed to be implemented onboard the vehicle with support of model compression techniques to eliminate latency caused by video wireless transmission and heavy computation burden on traditional teleoperation platforms.

1 Introduction

Unmanned Aerial Vehicles (UAVs) are maturing as a multi-disciplinary technology platform, their inherent aerial capabilities have been further developed and augmented, along with the advancement of powerful and robust perception techniques for aerial vehicles, UAVs are surpassing the traditional role as passive aerial observation platforms and able to actively interact with the environments to address aerial manipulation tasks in hard-to-reach or dangerous places. The tasks currently solved by aerial manipulators range from grasping [1], fetching [2], writing, peg-in-hole [3], to transporting arbitrary objects. However, all of these tasks are at an entry level and require only very basic tracking and interaction between UAV and the environments.

Real engineering problems in the daily life often involve sophisticated motion, contact and force control, also real-time precision target identification and tracking for environment perception. But most of these issues still remain hardly achievable under the traditional design, sensing and control of UAV systems, for example: inspection and failure detection of infrastructure like bridges or manufacturing plants, physical interactions through tools like grinding, welding, drilling, for construction projects or maintenance tasks in dangerous or harmful places [4].

In the aspect of aerial maneuverability, independent forces and torques must be exerted in certain degrees of freedom(DoF) to the environments to achieve a successful aerial manipulation operation, while the rest of DoFs of the end-effector of an aerial teleoperation platform should be accurately position-controlled [5]. However, traditional quadcopters are highly underactuated [6, 7], which means the translational motions on the horizontal plane are tightly coupled with rotational motions. Therefore, accurate position-level control and forces/torques control are impossible to be implemented simultaneously. In comparison to the traditional multirotor approaches, we can resolve the underactuation issue by proposing a novel tilting-rotor multirotor with an elegant and concise structure design, which is capable of generating 5-DoF thrust forces and torques, and completely avoiding inefficient force cancellation between rotors.

For the target identification and tracking of UAVs navigating in a large open space, a target detection-tracking approach has to be followed. Early solutions to object detection tasks depended on traditional machine learning methods, i.e., feature-based manual methods. The difficulty with the traditional approaches is that it is necessary to choose which features are important in each given image. As the number of classes to classify increases, feature extraction becomes more and more cumbersome [8]. Moreover, most of them are verified in low and medium density images and they usually need to be changed according to the specific situations [9], thus not suitable for applications in unstructured environments where illumination variations, partial occlusions, background clutter and shape deformation would occur. A deep learning model is trained based on the given data, and can automatically work out the most descriptive and salient features with respect to each specific class of object [10]. It has been demonstrated

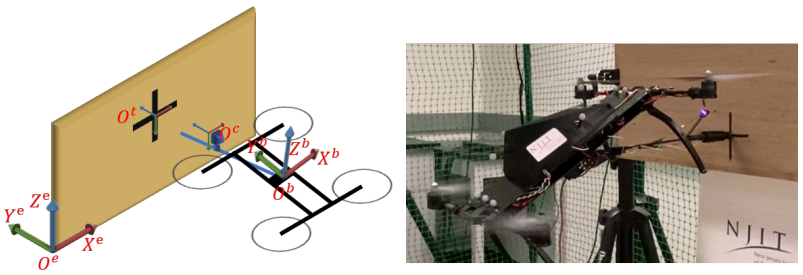


Figure 1: Prototype of the novel UAV cyber-physical system built in the Assistive and Intelligent Robotics Laboratory at NJIT: image stream from eye-in-hand camera is fed into the vision-based target identification and tracking module; target position and orientation are sent to the impedance control law for motion/force control.

that in many object detection applications, deep learning performs far better than traditional algorithms [11]. In this work, an onboard vision-based real-time guidance scheme is developed by dividing the identification and tracking problem into three parts: 1)Detect a target with a real-time YOLO v3 object detector, 2)2D transformation tracking using KLT tracker, 3)Extract position of the point of interest with morphological image processing method.

2 System Architecture And Mechanical Design

This section is to demonstrate the architecture and mechanical design of the developed tilting-rotor UAV, furthermore, to clarify the ability of independent forces/torques control, and independent position control in a mechanical design view. In addition, advantages of the UAV design are elaborated compared with previous aerial teleoperation platforms.

To perform aerial manipulation with traditional coplanar multirotors like quadcopter, hexacopter, and octocopter, aerial robots are equipped with a n-DoF robtic arm [12]. However, this solution comes with severe drawbacks. Firstly, a robotic arm strongly decreases the payload and flight time of an aerial vehicle due to its weight. Secondly, the system is much more complex mechanically, thus, it requires more maintenance and repairing efforts. Besides, lateral/longitudinal forces in body frame, which cannot be provided by the aerial platform itself, have to be generated through the dynamical/inertial coupling between the arm and the aerial robot, which is extremely hard to achieve in real-world conditions.

To solve both the underactuation problem and issues associated with robotic arms at once, many researchers have developed multirotor UAVs with fully-actuated aerial mobility. These designs use more than 6 rotors facing different directions [3], or include one/two extra servomotor for each rotor [13], to achieve 6-DoF actuation. However, the inclusion of a large number of motors or servo parts makes the designs mechanically more complicated, heavier, and the inefficient internal force cancellation can hardly be avoided [14].

The novel tilting-rotor multirotor UAV introduced in this work employees only two additional servomotors compared to traditional quadcopters, to actuate the two pairs of rotors on the vehicle. Specifically, the two pairs of rotors are mounted on two independently-actuated arms placed at both sides of the vehicle in an “H” configuration, as shown in Figure 1. Each arm is driven by a single servomotor, and carries two rotors on its two ends. Hence, the 5-DoF forces and torques can be generated by tilting the two arms.

3 Motion/Force Control And Real-time Target Tracking

This section presents the motion/force controller for position tracking and contact force regulation and the vision-based target identification and tracking system, these two modules are integrated together on the proposed UAV cyber-physical system for performing precision aerial manipulation tasks in unstructured environments, as shown in Figure 2.

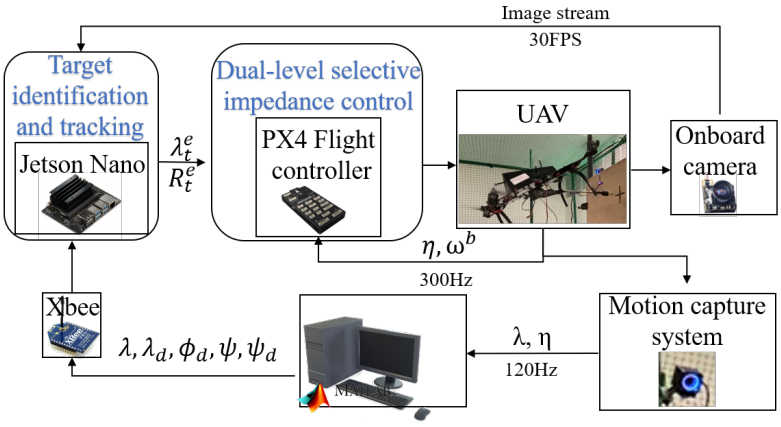


Figure 2: An overview of the UAV aerial manipulation system: the target identification and tracking algorithm is running onboard on the Jetson nano micro-controller to capture the target position and orientation in real-time; the dual-level impedance control algorithm is running on the PX4 flight controller at $300Hz$.

3.1 Motion/Force Control.

A dual-level cascaded control architecture is deployed for flight control. The low-level attitude controller uses a PID controller to generate updates of the required body-frame torques τ^b to accurately regulate the attitude angles, while a selective impedance controller is implemented at high level to indirectly regulate the contact force with environments F_p^b and track the target position captured by the vision-based target identification and tracking module. Finally, the five body-frame force and torque inputs will be used to generate the propeller thrust force commands and the tilt angle values to be fed into the rotors and servos respectively. In the fact that the PID control method has been widely applied to various commercial UAVs for the low-level attitude regulation, and has been developed as a standard approach with guaranteed stability and robustness verified in real-world applications, this work will focus on the introduction of high-level impedance controller for motion/force control of the proposed novel UAV.

Firstly, the total thrust force generated by the vehicle can be expressed in body frame as $F_p^b = [0 \;\; F_{py}^b \;\; F_{pz}^b]^T$. Defining F_1, F_2, F_3, F_4 as the absolute values of the thrust forces of the four propellers and α, β as the front and rear tilting angles, the F_{py}^b and F_{pz}^b can be expressed as:

$$\begin{aligned} F_{py}^b &= (F_1+F_4)S_\alpha + (F_2+F_3)S_\beta, \\ F_{pz}^b &= (F_1+F_4)C_\alpha + (F_2+F_3)C_\beta. \end{aligned} \tag{1}$$

In the same way, the total torque is expressed as $\tau^b = \begin{bmatrix}\tau_x^b & \tau_y^b & \tau_z^b\end{bmatrix}^T$. Defining l_1 as the half distance between the two rotors on each tilting axis, l_2 as the half distance between the two tilting axes, the components of τ_p^b can be expressed in body frame as

$$\begin{aligned} \tau_x^b &= (F_1+F_4)l_2C_\alpha - (F_2+F_3)l_2C_\beta, \\ \tau_y^b &= (F_1-F_4)(C_\alpha l_1 + S_\alpha k) + (F_2-F_3)(C_\beta l_1 - S_\beta k), \\ \tau_z^b &= (F_4-F_1)(S_\alpha l_1 - C_\alpha k) + (F_3-F_2)(S_\beta l_1 + C_\beta k). \end{aligned} \tag{2}$$

It's defined that $\lambda^t = [x^t \;\; y^t \;\; z^t]^T$ is the coordinate vector of vehicle's position O_b expressed in the target frame whose origin coincides with the target point and Z^t axis is perpendicular to the workpiece surface.

According to the Newton–Euler equations, differentiating λ^t twice, yields the following translation dynamic model

$$M\ddot{\lambda}^t = F_p^t + F_g^t + F_c^t, \tag{3}$$

where F_p^t, F_g^t, F_c^t are the total thrust force, gravity force, and contact force represented in target frame, respectively.

The selective impedance controller is designed as

$$F_p^t = -F_g^t - F_c^t + M\ddot{\lambda}_d^t - C\dot{e}_\lambda^t - Ke_\lambda^t, \tag{4}$$

where $e_\lambda^t = \lambda^t - \lambda_d^t$ is the position tracking error in target frame, λ_d^t is the desired position, $C = diag(C_x, C_y, C_z)$ and $K = diag(K_x, K_y, K_z)$ are diagonal matrices consisting of all the virtual damping and spring parameters along X^t, Y^t, Z^t. With the proposed impedance controller design, the following impedance model relating the position tracking error to the contact force can be derived:

$$M\ddot{e}_\lambda^t + C\dot{e}_\lambda^t + Ke_\lambda^t = F_c^t. \tag{5}$$

The damping and spring parameters for different axes can be tuned according to the desired hybrid motion/force control objective.

After obtaining the target frame thrust force F_p^t, the body frame thrust force can be calculated as $F_p^b = R_t^b F_p^t$, then F_{py}^b, F_{pz}^b can be derived directly. Finally, by setting $\alpha = \beta$ and solving the five equations in (1) and (2), the four individual rotor thrust forces and the tilt angle can be obtained.

3.2 Real-time Target Tracking.

A Deep Neutral Network (DNN) Based Objection Detection.

The most important task for the UAV target tracking mission is real-time object detection. The DNN-based object detection models are first trained on labeled images and are then used to predict the targets' bounding box and classification. The state-of-the-art DNN based object detection models can be divided into two classes: region proposal network (RPN) based detection method and single-stage detection method.

RPN based detection method. The RPN-based detection method requires one stage to extract the region of interest (RoI), and another stage to predict the classification result and bounding box. The famous RPN based methods include SPP network [15], R-CNN [16] and its' variant Faster R-CNN [17]. Although the R-CNN gains significant improvement in its' performance in recent years, its' inference speed is still limited by the high computation burden of the region proposal stage. Thus, R-CNN not suitable for the UAV's real-time detection application.

Single stage detection method. The single-stage detection eliminates the need for the region proposal stage and proposes an end-to-end structure to predict bounding boxes and class probabilities by a single evaluation step. The typical single-stage detection networks include the You Only Look Once (YOLO) [18] network and it's variants YOLOv2 [19], YOLOv3 [20], and YOLOv4 [21]. The variants of YOLO have improvement on both parameter size and prediction accuracy. Single-Short MultiBox Detector (SSD) [22] is another type of single-stage detection method. It features its' high prediction speed, but it has lower accuracy and is limited to a fixed number of bounding boxes. The challenge of single-stage detection lies in improving the inference speed while retaining good accuracy. For our UAV real-time objection detection application, YOLO and its' variants are more suitable than other DNNs.

B Inference Acceleration for Real-time Detection.

In order to achieve accelerating the DNN inference speed, model compression techniques are adopted to reduce the DNN model's parameter size. The major model compression techniques include weight pruning [23] and weight quantization [24]. There are unstructured pruning methods and structured pruning methods for the weight pruning technique. The unstructured pruning technique [25] prunes out the weight matrix in an irregular pattern and can achieve a higher compression ratio without much accuracy loss. However, due to its irregular memory access pattern, the unstructured pruning can hardly be accelerated on most of the hardware platforms. The structured pruning technique [26, 27, 28] constrains the weight matrix to be pruned in a structured and hardware-friendly pattern. For example, block-circulant matrices [24, 29, 30] can be used for weight representation after pruning. The structured pruning-based hardware implementation achieves better performance due to the higher parallelism achievable by regular memory access patterns and reduced computation burden.

For most of the objection detection tasks, the convolution layer is the central part of the networks. Fast Winograd algorithm and fast Fourier transform (FFT) algorithm can be used to accelerate the convolution operations and can

also accelerate the CNN models with small filters [31]. For example, 3×3 convolution layers make up 83% of the YOLOv4's weights and 81% of the YOLOv4's computations. Those algorithms bring the computation complexity of convolution operation down from $O(n^2)$ to $O(n \log n)$.

In order to reduce the communication latency and data transferring latency for UAV applications, object detection tasks need to be deployed on the onboard acceleration platform. Researchers are using the NVIDIA Jetson TX series embedded GPU platform to accelerate YOLO architecture for real-time object detection [32, 33, 34], and a 20 FPS detection speed is achieved in [34]. FPGA platform also gains popularity on edge computing application, [35] implemented YOLO network on Xilinx ADM-7V3 FPGA and achieved a 314.2 FPS detection speed. Simultaneously, the energy efficiency is seven times higher than that of the Jetson TX2 platform.

4 Conclusion

While UAVs have demonstrated great potential to be employed as teleoperation paltforms for aerial manipulation, the underactuation nature of traditional multirotors, and the latency caused by data transferring and heavy computation burden on traditional aerial teleoperation platforms still hinder the application of UAVs in aerial manipulation. To overcome these problems, we proposed a novel tilting-rotor UAV to cope with the underactuation issue, and an onboard real-time target identification and tracking approach to eliminate latency in the system. Based on the 5-DoF tilting-rotor UAV, a selective impedance control law is designed for motion/force control, which enables the UAV to track the target object whose position and orientation are captured by the onboard vision-based identification and tracking module, and to regulate contact forces between the end-effector and the environments.

References

[1] J. Thomas, G. Loianno, J. Polin, K. Sreenath, and V. Kumar, "Toward autonomous avian-inspired grasping for micro aerial vehicles," *Bioinspiration & biomimetics*, vol. 9, no. 2, p. 025010, 2014.

[2] S. Kim, H. Seo, S. Choi, and H. J. Kim, "Vision-guided aerial manipulation using a multirotor with a robotic arm," *IEEE/ASME Transactions On Mechatronics*, vol. 21, no. 4, pp. 1912–1923, 2016.

[3] S. Park, J. Lee, J. Ahn, M. Kim, J. Her, G.-H. Yang, and D. Lee, "Odar: Aerial manipulation platform enabling omnidirectional wrench generation," *IEEE/ASME Transactions on mechatronics*, vol. 23, no. 4, pp. 1907–1918, 2018.

[4] C. Ding, L. Lu, C. Wang, and C. Ding, "Design, sensing, and control of a novel uav platform for aerial drilling and screwing," *IEEE Robotics and Automation Letters*, pp. 1–1, 2021.

[5] C. Ding and L. Lu, "A tilting-rotor unmanned aerial vehicle for enhanced aerial locomotion and manipulation capabilities: Design, control, and applications," *IEEE/ASME Transactions on Mechatronics*, pp. 1–1, 2020.

[6] Y. Yu and X. Ding, "A global tracking controller for underactuated aerial vehicles: design, analysis, and experimental tests on quadrotor," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 5, pp. 2499–2511, 2016.

[7] C. Ding, L. Lu, C. Wang, and J. Li, "6-dof automated flight testing using a humanoid robot arm," in *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2018, pp. 217–222.

[8] J. Walsh, N. O'Mahony, S. Campbell, A. Carvalho, L. Krpalkova, G. Velasco-Hernandez, S. Harapanahalli, and D. Riordan, "Deep learning vs. traditional computer vision," in *Computer Vision Conference (CVC)*, 2019.

[9] Q. Zhang, Y. Liu, C. Gong, Y. Chen, and H. Yu, "Applications of deep learning for dense scenes analysis in agriculture: A review," *Sensors*, vol. 20, no. 5, p. 1520, 2020.

[10] H. F. Nweke, Y. W. Teh, M. A. Al-Garadi, and U. R. Alo, "Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges," *Expert Systems with Applications*, vol. 105, pp. 233–261, 2018.

[11] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, 2013.

[12] C. Korpela, M. Orsag, T. Danko, and P. Oh, "Insertion tasks using an aerial manipulator," in *2014 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, 2014, pp. 1–6.

[13] C. Ding, L. Lu, and C. Wang, "Energy-efficient adaptive robust control of vector thrust uavs with unknown inertia parameters," in *Dynamic Systems and Control Conference*, vol. 51913. American Society of Mechanical Engineers, 2018, p. V003T36A005.

[14] C. Ding, L. Lu, C. Wang, and B. Ouyang, "Modeling and control of fully actuated vector thrust unmanned aerial vehicles," in *Proceedings of the International Symposium on Flexible Automation 2018 International Symposium on Flexible Automation*. The Institute of Systems, Control and Information Engineers, 2018, pp. 451–458.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[16] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.

[17] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *arXiv preprint arXiv:1506.01497*, 2015.

[18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[19] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[20] ——, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[21] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[22] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*. Springer, 2016, pp. 21–37.

[23] B. Li, S. Pandey, H. Fang, Y. Lyv, J. Li, J. Chen, M. Xie, L. Wan, H. Liu, and C. Ding, "Ftrans: energy-efficient acceleration of transformers using fpga," in *Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design*, 2020, pp. 175–180.

[24] C. Ding, S. Liao, Y. Wang, Z. Li, N. Liu, Y. Zhuo, C. Wang, X. Qian, Y. Bai, G. Yuan *et al.*, "Circnn: accelerating and compressing deep neural networks using block-circulant weight matrices," in *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture*, 2017, pp. 395–408.

[25] T. Geng, T. Wang, C. Wu, C. Yang, W. Wu, A. Li, and M. C. Herbordt, "O3bnn: An out-of-order architecture for high-performance binarized neural network inference with fine-grained pruning," in *Proceedings of the ACM International Conference on Supercomputing*, 2019, pp. 461–472.

[26] Y. Cai, H. Li, G. Yuan, W. Niu, Y. Li, X. Tang, B. Ren, and Y. Wang, "Yolobile: Real-time object detection on mobile devices via compression-compilation co-design," *arXiv preprint arXiv:2009.05697*, 2020.

[27] B. Li, Z. Kong, T. Zhang, J. Li, Z. Li, H. Liu, and C. Ding, "Efficient transformer-based large scale language representations using hardware-friendly block structured pruning," *arXiv preprint arXiv:2009.08065*, 2020.

[28] G. Yuan, X. Ma, C. Ding, S. Lin, T. Zhang, Z. S. Jalali, Y. Zhao, L. Jiang, S. Soundarajan, and Y. Wang, "An ultra-efficient memristor-based dnn framework with structured weight pruning and quantization using admm," in *2019 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*. IEEE, 2019, pp. 1–6.

[29] L. Lu, Y. Liang, Q. Xiao, and S. Yan, "Evaluating fast algorithms for convolutional neural networks on fpgas," in *2017 IEEE 25th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*. IEEE, 2017, pp. 101–108.

[30] S. Liao, Z. Li, X. Lin, Q. Qiu, Y. Wang, and B. Yuan, "Energy-efficient, high-performance, highly-compressed deep neural network design using block-circulant matrices," in *2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2017, pp. 458–465.

[31] A. Lavin and S. Gray, "Fast algorithms for convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4013–4021.

[32] N. Bhandary, C. MacKay, A. Richards, J. Tong, and D. C. Anastasiu, "Robust classification of city roadway objects for traffic related applications," in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*. IEEE, 2017, pp. 1–6.

[33] E. Rudiawan, R. Analia, P. D. Sutopo, H. Soebakti *et al.*, "The deep learning development for real-time ball and goal detection of barelang-fc," in *2017 International Electronics Symposium on Engineering Technology and Applications (IES-ETA)*. IEEE, 2017, pp. 146–151.

[34] N. Smolyanskiy, A. Kamenev, J. Smith, and S. Birchfield, "Toward low-flying autonomous mav trail navigation using deep neural networks for environmental awareness," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 4241–4247.

[35] C. Ding, S. Wang, N. Liu, K. Xu, Y. Wang, and Y. Liang, "Req-yolo: A resource-aware, efficient quantization framework for object detection on fpgas," in *Proceedings of the 2019 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, 2019, pp. 33–42.

Resource Aware Learning-based Optimal Control of Cyber-Physical Systems

Avimanyu Sahoo[1], Vignesh Narayanan[2], and Jagannathan Sarangapani[3]
[1]Oklahoma State University, Stillwater, OK
[2] Washington University, St. Louis, MO
[3] Missouri University of Science and Technology, Rolla, MO

Abstract

Cyber-physical systems (CPS) are composed of the cyber components, such as sensing and information exchange infrastructures, and the physical components, such as actuation mechanisms and the plant, both of which are tightly integrated with each other. Owing to the advent of the internet-of-things and the advances in data science, CPS are ubiquitous in critical infrastructures such as smart energy systems, smart cities, automotive systems, medical prosthetic, and wearable devices. One of the objectives of a cyber-physical control system involves efficiently handling computational and communication constraints associated with the cyber components to reduce operational costs of the CPS. In this work, we present a resource aware optimal control scheme for efficiently controlling a CPS in a resource constrained setting. Specifically, we consider the problem of saving computational and communication resources when time-varying delays, packet losses, and bandwidth limitations are introduced via the cyber components of a CPS, which may jeopardize the stability of the physical system. To this end, we present a stochastic optimal regulator using adaptive dynamic programming and Q-learning techniques with event-sampled state and input vector. We design the tuning laws to update the parameters of the Q-function at the event-sampling instants and the events to adaptively determine the sampling and transmission instants. We show that the designed (events) sampling conditions not only warrant system stability but also save communication costs and facilitates the convergence of the Q-function parameters. We derive sufficient conditions that ensure asymptotic stability in the mean of the closed-loop system and demonstrate that the inter-sample times are non-trivial. Finally, we include numerical examples to substantiate the analytical design.

1 Introduction

Cyber-physical systems (CPS) [1] are composed of tightly integrated cyber and physical system components forming a sophisticated large-scale dynamic infrastructure. The rapid growth in CPS, seen in recent years, is triggered by the massive reduction in the need for hard wired connections between different components of a large-scale plant and the advances in communication and data sciences, making it more flexible for building distributed architectures. Common examples of such CPS include power transmission and distribution networks, water networks, large-scale robotic systems, underwater networks for oceanographic data collection, offshore exploration, disaster prevention, smart city and grid, and tactical surveillance systems. While offering flexibility, the CPS introduces various challenges, especially in ensuring its operability, efficiency and security, due to its many interconnected dynamical components. For example, the communication network in a CPS architecture have inherent constraints or imperfections, which introduces delays, packet losses, and link failure with bandwidth limitations. These network artifacts then propagate and affect the performance of the physical system components of the CPS. Hence, from a control-theoretical viewpoint, development of advanced feedback control approaches to retain the stability of the closed-loop system [2] along with optimal performance in the presence of such challenges is crucial. In the recent past, an ample amount of research has been carried out for studying the stability of networked control systems (NCS) [2, 3, 4], a control system closed by a communication network, in the presence of the aforementioned network constraints.

In general, in a CPS, a digital packet switched communication network is employed to transmit the feedback data between the system and the controller or between different subsystems. A fixed transmission frequency or interval is utilized to facilitate reliable information exchange among the different components of a control system as per the Shannon's sampling frequency [2, 3, 4, 5]. Thus, the feedback information is transmitted from the sensors to controller and the control actions from the controller to actuation mechanism, typically, at a predetermined transmission rate irrespective of the dynamic operating conditions of the system. However, when the network is constrained by limited bandwidth, such periodic transmissions of the signals poses a great challenge and increases operational costs. To alleviate this problem, an alternate framework, referred to as the event-triggered control [6, 7, 8], was proposed in the literature. The key rationale behind the event-triggered control framework is to dynamically decide the time

instants for transmitting signals between sensors, controllers, and actuators in the feedback control loop without sacrificing the stability and performance of the system. In the past decade, there have been significant efforts in the development of different event-triggered control schemes for linear and nonlinear systems [6, 7] in the presence of delays and packet losses [8] when an accurate model of the system is available [9] or when the system models are uncertain/unknown [10].

On the other hand, optimal control theory [11] offers tools for designing control policies for physical systems that help minimize the cost associated with a given control objective defined via a suitable performance index. Moreover, reinforcement learning (RL) techniques such as adaptive dynamic programming [12] have been investigated extensively in the literature [13, 14, 15] to synthesize forward-in time and online near optimal solution to an optimal control problem. These ADP/RL approaches integrates tools from adaptive control theory [16], optimal control [11], and RL [12] to obtain data-driven (approximate) solutions to optimal control problems when the system models are uncertain/unknown. Among the many learning ADP approaches that are available in the literature, model-free Q-learning [17] and actor-critic neural networks [12] for linear and nonlinear systems, respectively, are the most commonly used. While these approaches can be implemented even when the system model is unknown, typically, they perform a computationally intensive value/policy iterations to update the parameters in the learning mechanism. The authors in [3, 4] presented a time-based Q-learning approach to reduce the required iterations for learning by using the time history of the temporal difference (TD) or Bellman error to estimate the Q-function. In all the aforementioned schemes [3, 4], the control system operated under a fixed sampling/transmission rate to sample and transmit information in the control feedback loop, necessitating a large network bandwidth. Also, such a periodic communication protocol aggravates, and often induces, the network losses by increasing the delays and packet drop-outs.

As a result, event-sampled optimal control design methodologies with known and unknown dynamics are critically important to handle these challenges. In this context, event-driven ADP approaches have been proposed in the literature for linear, nonlinear, and interconnected dynamical systems [10, 18]. Furthermore, the design of a stochastic optimal controller using ADP for nonlinear NCS with uncertain dynamics was presented in [19]. To accelerate the learning, more recently, a hybrid Q-learning approach for linear NCS, which embeds time-varying event-dependent iterations in the learning mechanism was also introduced [20, 21]. Although, the hybrid approach accelerate the learning it increases the computation.

In this paper, we focus on an optimal adaptive regulator, introduced in [22], via Q-learning for an NCS that is modeled as an uncertain continuous-time system governed by linear dynamics in the presence of the packet losses and time-varying delays with event sampled input and state vector. In particular, we investigate the Q-learning-based control methodology for an NCS wherein the system state vector and control inputs are available to the controller and the actuator, respectively, at event sampled instants to design stochastic optimal regulators without relying on the knowledge of an accurate model representing the system dynamics. In contrast to traditional event-triggered control [7] or the Lyapunov based event-triggering scheme in [20], here, the events are designed using an adaptive event sampling condition which facilitates the learning mechanism to estimate the Q-function parameters while retaining the advantages of the traditional event-triggered control methodology in terms of its superior resource saving features. We also derive sufficient conditions to ensure a non-trivial inter-sample times and closed loop stability of the control system.

The remainder of the paper is organized as follows. In Section 2, we give a brief background of NCS and formulate the problem considered here. The event-based Q-learning is presented in Section 3 followed by simulation results in Section 4. We include conclusions and future outlook in Section 5.

2 NCS Reformulation and Problem Statement

2.1 NCS Reformulation and Background

We consider an NCS represented by linear time-invariant (LTI) control system that evolves continuously with time. The feedback loop is assumed to be closed via a communication channel, which induces time-varying delays and packet losses. The dynamics of such an NCS is represented by

$$\dot{x}(t) = Ax(t) + \alpha(t)Bu(t-\tau(t)), \;\; x(0) = x_0, \tag{6}$$

where $u(t) \in \Re^m$ and $x(t) \in \Re^n$ are the control input and system state vectors, respectively, and the matrices $A \in \Re^{n\times n}$ and $B \in \Re^{n\times m}$, model the time-invariant system and input dynamics, respectively. We assume that these matrices are uncertain/not accurately known. The random time delay $\tau(t) = \tau_{sc}(t) + \tau_{ca}(t)$, is composed of mutually independent delays $\tau_{sc}(t)$ and $\tau_{ca}(t)$ from the sensors to the controller (S-C) and the controller to the actuator (C-A) channels, respectively. The random packet losses introduced by the communication channel is modeled via an indicator $\alpha(t) = \alpha_{sc}(t)\alpha_{ca}(t)$ defined as

$$\alpha(t) = \begin{cases} 1, & \text{packet received at } t; \\ 0, & \text{packet lost at } t. \end{cases} \tag{7}$$

where the packet loss indicators $\alpha_{ca}(t)$ and $\alpha_{sc}(t)$ corresponds to the C-A and the S-C channel. It is assumed that the random packet losses in both the channels are independent. Additionally, the design and analysis presented in this paper involve the following assumptions.

Assumption 1. *(i) The matrix pair (A,B) renders the system in (6) controllable. The order of the system is known, and further, the states are measurable; and (ii) The initial conditions of the system states are deterministic and the matrix modeling the input dynamics B is bounded and satisfies $\|B\| \leq B_{max}$, with a known constant bound B_{max}.*

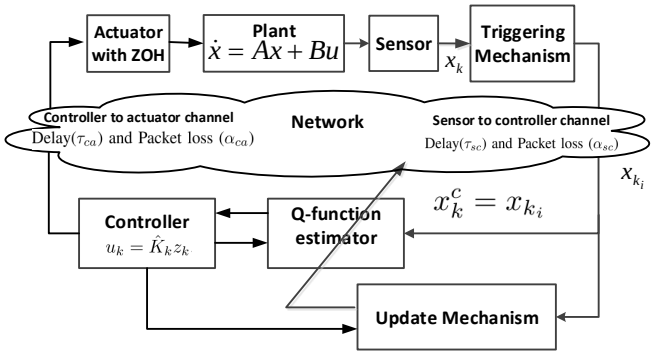


Figure 1: Proposed event-triggered NCS.

The NCS control architecture considered in this work is represented as a block diagram in Fig. 1. Note that because the control feedback loop is closed via a packet switch communication network, we naturally consider the implementation of the controller in a sampled data frame work. In this context, the system states are sampled by the sensor at time instants kT_s, where $k \in \mathbb{N}$ with a fixed sampling period T_s. These measurements and the control input are transmitted between different components in the control loop at event-triggering instants (to be designed), and between successive events they remain constant (held via a zero-order hold mechanism). To make the design tractable in the presence of network losses, we assume that the following standard properties of the communication network and the system hold true [5].

Assumption 2. *[5] (i) The time varying random delays at the k^{th} time instant, introduced in the communication channel, are bounded and satisfy $\tau_{sc}^k \leq \Delta_s$ and $\tau_{ca}^k \leq dT_s$. The skew between the sensor ad the controller sampling instants is denoted by $\Delta_s < T_s$ and the delay bound $d > 0$ is an integer. Further, the time varying delays τ_{ca}^k and τ_{sc}^k are independent and the distribution is known; (ii) The random packet losses α_{ca}^k and α_{sc}^k are assumed to be i.i.d processes following the Bernoulli distribution with $\mathbb{P}(\alpha^k = 1) = \mathbb{P}(\alpha_{sc}^k = 1)\mathbb{P}(\alpha_{ca}^k = 1) = \alpha_p$, where the probability operator is denoted by $\mathbb{P}(\cdot)$.*

As the C-A channel delay τ_{ca}^k is bounded by the positive integer d (Assumption 2 (i)), there may be a maximum of $d+1$ control inputs available at the actuator in the interval $[kT_s, (k+1)T_s]$, without any packet loss, i.e., $\alpha_k = 1$. The latest control input is applied to the system and held using a the ZOH till the next is available at time $kT_s + t_l^k$ where t_l^k, $l = 0,1,2,\cdots,d$ is the time after kT_s satisfying $t_l^k > t_{l+1}^k$ and $t_{-1}^k = T_s$ and $T_d^k = 0$, as shown in Fig. 2.

By integrating the system dynamics in (6) in the time interval $[k,\ k+1]$ [5], the discrete time dynamics can be expressed as

$$x_{k+1} = A^s x_k + \sum_{l=0}^{d} \alpha_{k-l} B_l^k u_{k-l}, \tag{8}$$

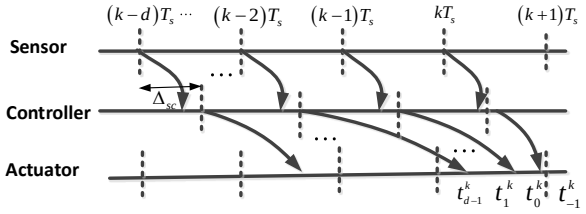


Figure 2: Feedback data transmission under time-varying delay.

where $x_k = x(kT_s)$, $A^s = e^{AT_s}$, $B^k_l = \int_{t^k_l}^{t^k_{l-1}} e^{A(T_s-\tau)} d\tau B$ with $t^k_{-1} = T_s$ and $t^k_d = 0$.

Defining the augmented state $z_k = [x^T_k \; u^T_{k-1} \; \cdots \; u^T_{k-d}]^T \in \Re^{n+dm}$, the system in (8) can be reformulated as

$$z_{k+1} = A^{\tau,\alpha}_k z_k + B^{\tau,\alpha}_k u_k \tag{9}$$

where $A^{\tau,\alpha}_k \in \Re^{(n+dm)\times(n+dm)}$, $B^{\tau,\alpha}_k \in \Re^{(n+dm)\times m}$, defined as $A^{\tau,\alpha}_k = \begin{bmatrix} A^s & \alpha_{k-1}B^k_1 & \cdots & \alpha_{k-l}B^k_l & \cdots & \alpha_{k-\bar{d}}B^k_{\bar{d}} \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & I_m & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & I_m & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & \cdots & I_m & 0 \end{bmatrix}$

and $B^{\tau,\alpha}_k = \begin{bmatrix} \alpha_k B^k_0 \\ I_m \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$, are stochastic time varying system matrices. The augmented system matrices $A^{\tau,\alpha}_k$ and $B^{\tau,\alpha}_k$ are stochastic and time-varying. The integer d is selected such that the pair $(A^{\tau,\alpha}_k, B^{\tau,\alpha}_k)$ [4] is controllable. From Assumption 2 (a) it is clear that $\|B_{z,k}\| \leq B_{z,\max}$ where $B_{z,\max} > 0$ a known constant.

The objective is design an optimal control policy by minimizing the infinite horizon stochastic performance index given by

$$J(x_0) = \mathop{\mathbb{E}}_{\tau,\alpha}\left\{\sum_{k=0}^{\infty}\left(x^T_k H x_k + u^T_k R u_k\right)\right\} \tag{10}$$

where the user defined penalty matrices $H \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ are positive definite. The operator $\mathop{\mathbb{E}}_{\tau,\alpha}\{\centerdot\}$ is the expected value of $(\cdot)$. The performance index for the augmented NCS (9) can be redefined as

$$J(z_k) = \mathop{\mathbb{E}}_{\tau,\alpha}\left\{\sum_{j=k}^{\infty} r(z_j,u_j)\,\middle|\, z_k\right\} \tag{11}$$

where $r(z_k,u_k) = z^T_k H_z z_k + u^T_k R_z u_k$ with $H_z = diag\{H, R/(d+1), \cdots, R/(d+1)\}$, $R_z = R/(d+1)$ are positive definite matrices. A stochastic Q-learning based scheme [4] with periodic state information can be employed to compute the optimal control policy and presented below for completeness.

The stochastic the Q-function or action dependent value function using the performance index (11) can be written as

$$Q(z_k,u_k) = \mathop{\mathbb{E}}_{\tau,\alpha}\{(r(z_k,u_k) + V^*(z_{k+1}))|z_k\} \tag{12}$$

where $Q(z_k,u_k)$ is the Q-function and $V^*(z_{k+1})$ is the optimal cost for $k+1$ onward. Since, the optimal value is quadratic, we have $V^*(z_{k+1}) = \mathop{\mathbb{E}}_{\tau,\alpha}\{z^T_{k+1} S_{k+1} z_{k+1}\}$. Substituting $V^*(z_{k+1})$ in (12), we have $Q(z_k,u_k) = \mathop{\mathbb{E}}_{\tau,\alpha}\{z^T_k P_z z_k + u^T_k R_z u_k + z^T_{k+1} S_{k+1} z_{k+1}\}$.

Along the system dynamics (9), the Q-function can be expressed in a compact form as

$$Q(z_k,u_k)=\underset{\tau,\alpha}{\mathbb{E}}\{[z_k^T \quad u_k^T]\bar{G}_k[z_k^T \quad u_k^T]^T\}=\underset{\tau,\alpha}{\mathbb{E}}\{w_k^T\bar{G}_k w_k\},\forall k \tag{13}$$

where $\bar{G}_k=\begin{bmatrix}\bar{G}_k^{zz} & \bar{G}_k^{zu}\\ \bar{G}_k^{uz} & \bar{G}_k^{uu}\end{bmatrix}=\begin{bmatrix}H_z+\underset{\tau,\alpha}{E}\{A_k^{\tau,\alpha^T}S_{k+1}A_k^{\tau,\alpha}\} & \underset{\tau,\alpha}{E}\{A_k^{\tau,\alpha^T}S_{k+1}B_k^{\tau,\alpha}\}\\ \underset{\tau,\alpha}{E}\{B_k^{\tau,\alpha^T}S_{k+1}A_k^{\tau,\alpha}\} & R_z+\underset{\tau,\alpha}{E}\{B_k^{\tau,\alpha^T}S_{k+1}B_k^{\tau,\alpha}\}\end{bmatrix}$ and $w_k=[z_k^T \quad u_k^T]^T\in\Re^{l_{mn}^1}$ with $n+(d+1)m=l_{mn}^1$.

The optimal Q-function $Q^*(z_k,u_k^*)$ is equals the optimal value function $V^*(z_k)$ when $u_k=u_k^*$, i.e., $\underset{\tau,\alpha}{\mathbb{E}}\{Q^*(z_k,u_k^*)\}=\underset{\tau,\alpha}{\mathbb{E}}\{V^*(z_k)\}$. Therefore, the optimal control policy with gain K_k^* can be written as

$$u_k^*=K_k^*z_k \tag{14}$$

where $K_k^*=(\bar{G}_k^{uu})^{-1}\bar{G}_k^{uz}$. By estimating the $\bar{G}_k$ matrix with the periodic state and input information the control policy can be computed. Traditionally, the (13) is represented in a parametric form and estimated using adaptive control technique using the following standard trivial assumption.

Assumption 3. *The Q-function parameters vary slowly with time. Further, $Q(z_k,u_k)$ can be represented in a linear in the unknown parameters form.*

However, to reduce the transmission and computation using the event-sampled control approach the NCS state and control policy is not available in a periodic manner. The problem of estimation and control in an event-sampled system frame work is presented next.

2.2 Problem Statement

In an event-sampled control framework for the NCS, the periodic sampled state is evaluated by a triggering mechanism at each k whether to transmit or not. Define the transmission/sampling instants $\{k_i\}_{i=0}^{\infty}$, a sub-sequence of $k\in\mathbb{N}$ with $k_0=0$. The trigger mechanism employs a state dependent condition (discussed later) to determine the transmission instants $k_i,\forall i=1,2,\cdots$ as shown in Fig. 1.

The NCS state at the skewed controller time instants is given by

$$x_k^c=x_{k_i},\ \ k_i\le k<k_{i+1},\ i=1,2,\cdots. \tag{15}$$

where x_k^c is the event-sampled state received at the controller.

Remark 1. *The event-sampling (triggering) mechanism shown in Fig. 1 equipped with a mirror Q-function estimator to evaluate the event condition (presented later). Since, the delay $\tau_{sc}^k<\Delta_s$, the NCS state information at the mirror estimator and the controller are same and both operates in synchronism.*

The state vector z_k^s at the event-sampling instants can be formed by storing and augmenting sampled states as $z_k^s=z_{k_i},\ k_i\le k<k_{i+1},\forall i=1,2,\cdots$, where $z_{k_i}=[x_{k_i}^T \quad u_{k_{i-1}}^T \quad u_{k_{i-2}}^T \quad \cdots \quad u_{k_{i-\bar{d}}}^T]^T,\forall i=1,2,\cdots$. The error introduced in the reformulated NCS state by the event-sampled transmission referred to as event-sampling state error is given by

$$e_{ET,k}=z_k-z_k^s,\ \ k_i\le k<k_{i+1},\forall i=1,2,\cdots. \tag{16}$$

Note that the error $e_{ET,k}$ resets to zero once the new state information is arrived at the controller.

The optimal control input (14) with event sampled state vector z_k^s introduces error in the control policy, which may lead to instability. Further, the event-sampled state error $e_{ET,k}$ also drives the Bellman error used for estimating the Q-function parameters, which may affect the Q-function parameter estimation accuracy. This requires a trade-off between the number of sampling instances and control performance. Alternatively, unlike the traditional event-triggering conditions [7], the sampling condition must be driven by the Q-function parameter estimation error. Further, to update the Q-function parameters with updated state information, which are available at the aperiodic

sampling instants only, the parameter update law must only be executed at the event-sampling instants. Although, it is clear that the frequency of parameter updates will reduce due to the event-based execution when compared to discrete time adaptive control approaches, it may affect the estimation accuracy. Since the event-sampling error orchestrates the stability and Q-function parameter estimation, we refer the proposed approach as event-driven adaptive dynamic programming (E-ADP).

3 E-ADP using Q-learning

In this section, we present the event-based estimation of the Q-function to compute the near optimal control gain using the aperiodic state information received and stored at the controller. An event-sampling (triggering) condition is introduced to determine the sampling and transmission instants, which is derived using the Lyapunov based stability analysis.

The Q-function (13) in the vector form can be represented as

$$Q(z_k,u_k) = \underset{\tau,\alpha}{\mathbb{E}}\{\bar{g}_k^T \xi_k\}, \forall k \tag{17}$$

where $\bar{g}_k = vec(\bar{G}_k) \in \Re^{l_g^1}$ with $l_g^1 = (l_{mn}^1)^2$, $vec(\cdot)$ is the vectorization operator on matrix, and $\xi_k = w_k \otimes w_k \in \Re^{l_g^1}$ is the quadratic polynomial regression vector with $\otimes$ denotes the Kronecker product.

The estimated value of the stochastic Q-function can be expressed as

$$\hat{Q}(z_k,u_k) = \underset{\tau,\alpha}{\mathbb{E}}\{w_k^T \hat{\bar{G}}_k w_k\} = \underset{\tau,\alpha}{\mathbb{E}}\{\hat{\bar{g}}_k^T \xi_k\} \tag{18}$$

where $\hat{\bar{G}}_k = \begin{bmatrix} \hat{\bar{G}}_k^{zz} & \hat{\bar{G}}_k^{zu} \\ \hat{\bar{G}}_k^{uz} & \hat{\bar{G}}_k^{uu} \end{bmatrix}$ is the estimates of $\bar{G}$ and $\hat{\bar{g}}_k \in \Re^{l_g^1}$ is the vector form of $\hat{\bar{G}}_k$.

By Bellman principle of optimality, using the parametric form (17), the value function satisfies $0 = \underset{\tau,\alpha}{E}\{r(z_k,u_k) + \bar{g}_k^T \Delta\xi_k\}$, where $\Delta\xi_k = \xi_{k+1} - \xi_k$. The Bellman or temporal difference (TD) error with the estimated Q-function (18) can be expressed as $e_{V,k} = \underset{\tau,\alpha}{E}\{r(z_k,u_k) + \hat{\bar{g}}_k^T \Delta\xi_k\}$.

The estimation of the Q-function in a parametric form (18) with event sampled state vector z_k^s and event-sampled control policy $u_{e,k}$ is rewritten as

$$\hat{Q}(z^s{}_k, u_{e,k}) = \underset{\tau,\alpha}{E}\{\hat{w}_k^T \hat{\bar{G}}_k \hat{w}_k\} = \underset{\tau,\alpha}{E}\{\hat{\bar{g}}_k^T \hat{\xi}_k\}, k_i \le k < k_{i+1},\ \forall i = 1,2,\cdots \tag{19}$$

where $\hat{\xi}_k = \hat{w}_k \otimes \hat{w}_k$ is the event sampled regression vector with $\hat{w}_k = [z_k^{s^T} \quad u_{e,k}^T]^{\mathbf{T}}$. The event sampled Bellman error by using (19) becomes

$$e_{V,k} = \underset{\tau,\alpha}{E}\{r(z^s{}_k, u_{e,k}) + \hat{\bar{g}}_k^T \Delta\hat{\xi}_k\}, k_i \le k < k_{i+1}, \forall i = 1,2,\cdots. \tag{20}$$

The Bellman error (20) at the event sampled instants k_i with updated state and control policy information (($e_{ET,k} = 0$)) becomes

$$e_{V,k} = \underset{\tau,\alpha}{E}\{r(z_k,u_k)\} + \hat{\bar{g}}_k^T \Delta\xi_k,\ k = k_i, \forall i = 1,2,\cdots. \tag{21}$$

An augmented Bellman error, to accelerate the parameter estimation, using the time history can be written as

$$E_{V,k} = \Pi_k + \hat{\bar{g}}_k^T Z_k,\ k = k_i, \forall i = 1,2,\cdots \tag{22}$$

where $\Pi_k = [\underset{\tau,\alpha}{E}\{r(z_{k_i},u_{k_i})\} \ \underset{\tau,\alpha}{E}\{r(z_{k_{i-1}},u_{k_{i-1}})\} \ \cdots \ \underset{\tau,\alpha}{E}\{r(z_{k_i-1-v},u_{k_{i-1-v}})\}] \in \Re^{1\times v}$ and $Z_k = [\Delta\xi_{k_i} \ \Delta\xi_{k_{i-1}} \ \cdots \ \Delta\xi_{k_{i-1-v}}]$, $Z_k \in \Re^{l^1{}_g \times v}$. Note that the convergence of $E_{V,k}$ to the origin guarantees the convergence of$e_{V,k}$. Further, it is observed heuristically that the length of the history v determines the convergence rate and can be determined by designer experience.

The Q-function parameter update law to drive the Bellman error to zero can be selected from the Lyapunov analysis as

$$\hat{\bar{g}}_k = \begin{cases} \hat{\bar{g}}_{k-1} - \left(\alpha_V Z_{k-1} E_{V,k-1}^T / \left\| I + Z_{k-1}^T Z_{k-1} \right\|\right), & k = k_i, \\ \hat{\bar{g}}_{k-1}, & k_{i-1} < k < k_i, \end{cases} \tag{23}$$

where $\alpha_V > 0$ is the adaptive gain parameter. The parameter update law (23) is executed only at the aperiodic sampling instants, which reduces the computation and communication when compared to a periodic counter parts.

Remark 2. *The computation of the regression matrix Z_{k_i-1} in (22) requires the augmented state z_{k_i} and z_{k_i-1}. The received state vectors, $x_{k_i}$and x_{k_i-1} at the controller is stored to form the augmented state vectors in future instants.*

Remark 3. *The persistency of the excitation (PE) condition [16] is necessary for the regression matrix Z_k to ensure the convergence of Q-function parameter estimation error $\tilde{\bar{g}}_k$ to the origin. A PE like condition can be ensured by adding an exploration noise to the control input [3].*

By using the estimated Q-function parameters, the event sampled estimated control policy can be expressed as

$$u_{e,k} = -\hat{K}_k z^s{}_k, \; k_i \leq k < k_{i+1}, \forall i = 1, 2, \cdots \tag{24}$$

where $\hat{K}_k = \left(\hat{\bar{G}}_k^{uu}\right)^{-1} \hat{\bar{G}}_k^{uz}$ with $\hat{\bar{G}}_k^{uz}$ and $\hat{\bar{G}}_k^{uu}$ defined in (14).

The closed-loop dynamics of the event sampled system by using (9) and (24) is given by

$$z_{k+1} = A_k^{\tau,\alpha} z_k - B_k^{\tau,\alpha} \hat{K}_k z_k + B_k^{\tau,\alpha} \hat{K}_k e_{ET,k}, \; k_i \leq k < k_{i+1}. \tag{25}$$

To ensure that the the estimated near-optimal control policy (24) retains the stability of the NCS, we introduce the following adaptive event-sampling (triggering) condition. The NCS states are transmitted when the condition

$$\|e_{ET,k}\| \leq \sigma_{ET,k} \|z_k\| \tag{26}$$

is violated where $\sigma_{ET,k} = \sqrt{\alpha(1-3\kappa)/(\varepsilon + 3B_{z,\max}^2 \left\|\hat{K}_k\right\|^2)}$ is the adaptive threshold coefficient , $0 < \alpha < 1$, and $\hat{K}_k$ is defined in (24). The constant κ satisfies $\left\| \underset{\tau,\alpha}{E} \{A_{z,k} z_k + B_{z,k} u_{p,k}^*\} \right\|^2 \leq \kappa \|z_k\|^2$ and $\varepsilon > 0$ is a small constant. Further, it was shown using Lyapunov stability analysis that $0 < \kappa < 1/3$ will ensure asymptotic stability in the mean of the event sampled system, presented in the next theorem.

Theorem 3.1. *Consider the system dynamics (9), Q-function estimator (18) and controller (24) with parameter update law (23). Let the Assumptions 1 and 2 hold, the regression vector, ξ_k, satisfies the PE condition, and $\hat{\bar{g}}_0 \in \Omega_g$. Given an initial admissible control policy $u_0 \in \Omega_u \subset \Re^m$ and adaptive gain parameter satisfying $0 < \alpha_V < 2$, the closed-loop event sampled system (26) is asymptotically stable in the mean as $k_i \to \infty$. In addition, the control input $u_{e,k} \to u_{e,k}^*$ as $k_i \to \infty$ or alternatively, $k \to \infty$.*

Remark 4. *Note that the Q-function parameters and the control policy are updated at the event-sampling instants and held during the inter-sample times. This may lead to situation where the Lyapunov function may not decrease monotonically for all time instants k [8]. Therefore, the asymptotic stability in the mean [23] for the closed-loop system (25) is shown by considering a single Lyapunov function candidate, L_k, for both $k = k_i$ and $k_i \leq k < k_{i+1}$, and the existence of a piecewise continuous function $h(k) \in \Re^+$, such that $h_k \geq L_k$ and $\lim_{k\to\infty} h_k = 0$, $k \in \mathbb{N}$ hold, as illustrated in Figure 3.*

The event sampling condition is evaluated at every time instants k with a fixed sensor sampling time T_s. Therefore, the minimum time between two consecutive event sampled instants is T_s. Alternatively, $\delta k_{\min} = \min_{i\in\mathbb{N}} \{k_{i+1} - k_i\} = 1$. A condition to achieve the non-trivial inter-sample times i.e., $\delta k_i = k_{i+1} - k_i > 1$, to save the communication and computational load, is presented next.

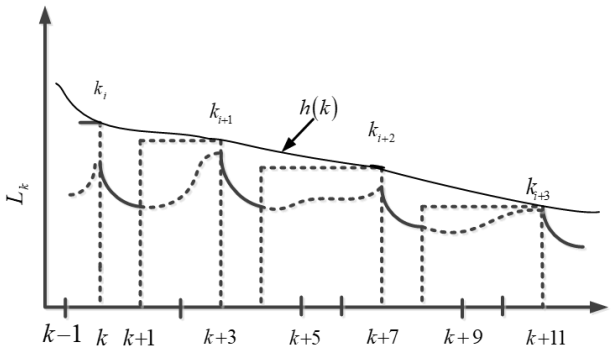


Figure 3: Lyapunov function upper bounded by a monotonically decreasing function during the event-sampling instants and inter-sample times.

Corollary 3.2. *Consider the NCS (9) with the Q-function estimator (18) and the state feedback controller (24). Then closed-loop event sampled system (25) is asymptotically stable in the mean with the event sampling condition given by*

$$\|e_{ET,k}\| \leq (\sigma_{ET,k}/(1+\sigma_{ET,k}))\,\|z^s{}_k\|\,. \tag{27}$$

We will use the results of the corollary to show the non-trivial triggering time and reduction in computation in the following theorem.

Theorem 3.3. *Given the hypothesis of Theorem 3.1 with the event sampling condition (27), the inter-sample times $\delta k_i = k_{i+1} - k_i$ implicitly defined by (27) satisfies $\delta k_i \geq \ln(1+(1/M_i)(F-1)\sigma_i)/\ln(F)$, $\forall i = 1,2,\cdots$, where $\sigma_i = \sigma_{ET,k_i}/(1+\sigma_{ET,k_i})$ for the i^{th} inter-sample time. Further, $F = \sqrt{\mu} + B_{z,\max}K^*_{\max}$ and $M_i = (\sqrt{\mu} + B_{z,\max}\|\tilde{K}_{k_i}\| + 1)$ with $K^*_{\max}$ is the maximum value of the optimal control gain matrix. Further the inter-sample times δk_i become non-trivial when $\sigma_i/M_i > 1$.*

Remark 5. *The function M_i and the threshold coefficient σ_{ET,k_i} depend on the control gain estimation error $\tilde{K}_k$ via the relation $\tilde{K}_k = K^*_k - \hat{K}_k$. Hence, the inter-sample times δk_i are a function of $\tilde{K}_k$ or $\tilde{\tilde{g}}_k$. It is clear that the convergence of $\tilde{\tilde{g}}_k$ close to zero, as proven Theorem 3.1, will satisfy the non-triviality condition. This further implies that the number of triggers will depend on the initial Q-function parameters and the adaptive learning gains.*

4 Simulation Results

In this section, we present the numerical results to validate analytical design using the batch reactor example. The linear continuous-time dynamics of the batch reactor is give as

$$\dot{x} = \begin{bmatrix} 1.38 & -0.2077 & 6.715 & -5.676 \\ -0.5814 & -4.29 & 0 & 0.675 \\ 1.067 & 4.273 & -6.654 & 5.893 \\ 0.048 & 4.273 & 1.343 & -2.104 \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ 5.679 & 0 \\ 1.136 & -3.146 \\ 1.136 & 0 \end{bmatrix} u. \tag{28}$$

The parameters used for the numerical experiment are presented next. Initial values of the augmented states are $z_0 = \begin{bmatrix} 0.2 & 1 & -3 & 0.5 \end{bmatrix}^T$ and initial values of the Q-function parameters $\hat{\tilde{g}}_0$ are chosen randomly in the interval $[0,\ 1]$. The delay bound $\bar{d} = 2$ with a mean value of 12.5 ms, as shown in Figure 3 (a). The packet loss α generated randomly from Bernoulli distribution with $p = 0.8$, as shown in Figure 3 (b). The periodic sampling time $T_s = 0.01$ sec and the event-based sampling is determined using (26). A penalty matrices of the cost function (11) were $P_z = 10^{-3}I_{8\times 8}$ and $R_z = 10^{-3}I_{2\times 2}$ where I is the identity matrix. The learning gains are $\alpha_V = 0.05$, $\Gamma = 0.99$, and $\kappa = 0.25$. A Monte Carlo simulation is run for 25 sec or 2500 sampled instants.

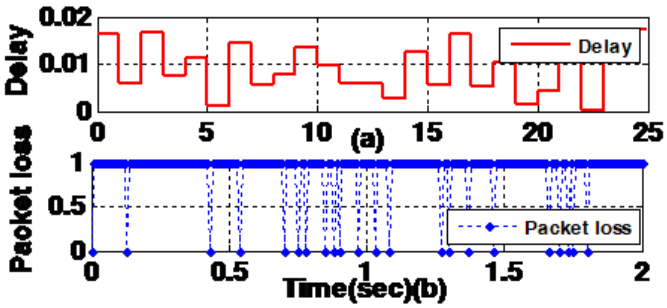


Figure 4: Distribution of (a) delays; and (b) packet losses.

The simulation results are compiled in Figures 4 to 6. The estimated optimal control policy, shown in 5 (b), under the influence of the random delays, packet losses and uncertain system dynamics, regulated the state vector to zero and shown in Figure 5 (a). The aperiodic Q-function parameter update laws guaranteed the convergence of the event-sampled Bellman error to zero as shown in Figure 5 (c). This implies that the estimated Q-function parameters converged to their optimal values before the system states converge to zero.

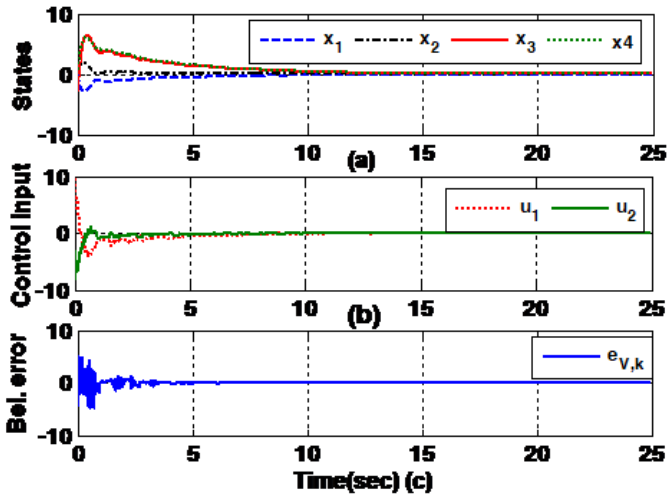


Figure 5: Convergence of: (a) closed-loop state vector; (b) event sampled optimal control policy; and (c) event sampled Bellman error

From the resource saving perspective, the reduction in computation and communication is clearly visible from the inter-event time plot in Figure 6 (b). The inter-sampling times are represented by the vertical lines and the minimum inter-sample-time observed during the simulation is 0.01 sec, which is equal to the sensor sampling time. It is to note that inter-sample times increased as the Q-function parameters converged close to the optimal values. Further, Figure 6 (a) shows the evolution of the the event-sampling state error and the threshold.

The mean value of the number of transmission/sampling during the simulation experiment was found to be 853 as shown in Figure 7 (a). This implies a significant reduction in transmission and computation. Figure 7 (b) depicts the bandwidth usage, assuming a packet size of 8 bit data, when compared with traditional periodic transmission schemes. It is clear that data rate is high during the initial estimation phase, which is closer to the traditional periodic transmission counter part. With the convergence of the Q-function parameter estimation errors the data rate also reduced significantly implying lesser bandwidth usage.

5 Conclusions and Future Outlook

In this letter, we proposed an event-sampled adaptive Q-learning to design an optimal controller for an NCS. The event-sampled controller under the time-varying delays and random packet losses regulated the system states to the origin with lesser number of transmission and computation. The event-sampling (triggering) condition was observed

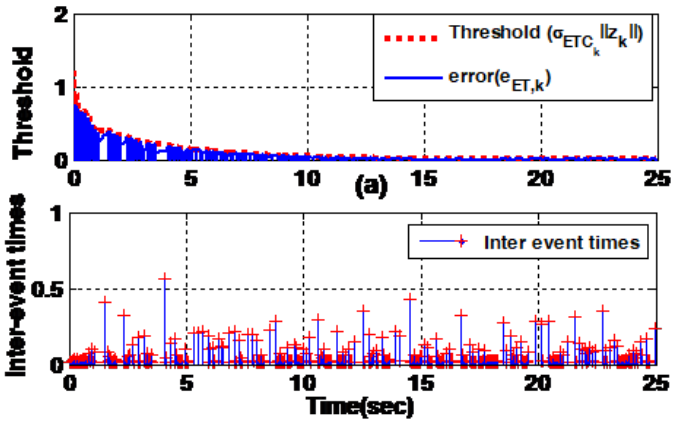


Figure 6: Evolution of (a) the threshold and event sampling error; and (b) inter-sample times.

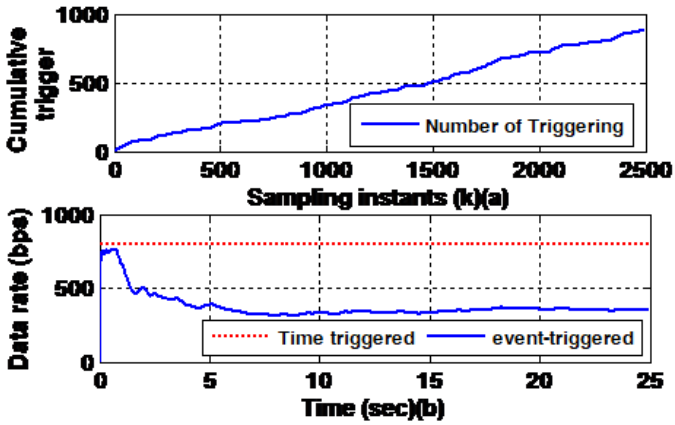


Figure 7: Comparison of the: (a) cumulative number of events with sampled instants; and (b) data rate between periodic and event-triggered system.

to be adaptive in nature and gets updated with the updated Q-function parameters. This generated the required number of transmissions for estimation of the Q-function. The aperiodic Q-function parameter tuning law regulated the state and parameter estimation errors to zero in the mean square. The numerical results with the batch reactor example corroborated the analytical designs from both the transmission and computation saving perspective. There are several other challenges one must account while developing the feedback control for CPS, such as adversarial attacks on the sensors, controllers, and communication network. Further, the communication becomes more challenging in a multi-agent CPS scenario and is an interesting area of future research.

References

[1] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *2008 The 28th International Conference on Distributed Computing Systems Workshops*. IEEE, 2008, pp. 495–500.

[2] G. C. Walsh, H. Ye, and L. G. Bushnell, "Stability analysis of networked control systems," *IEEE transactions on control systems technology*, vol. 10, no. 3, pp. 438–446, 2002.

[3] H. Xu, S. Jagannathan, and F. L. Lewis, "Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses," *Automatica*, vol. 48, no. 6, pp. 1017–1030, 2012.

[4] H. Xu, S. Jagannathan, and F. Lewis, "Stochastic optimal design for unknown linear discrete-time system zero-sum games in input-output form under communication constraints," *Asian Journal of Control*, vol. 16, no. 5, pp. 1263–1276, 2014.

[5] L.-W. Liou and A. Ray, "A stochastic regulator for integrated communication and control systems: Part i—formulation of control law," 1991.

[6] K. J. Åström and B. Bernhardsson, "Comparison of periodic and event based sampling for first-order stochastic systems," *IFAC Proceedings Volumes*, vol. 32, no. 2, pp. 5006–5011, 1999.

[7] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1680–1685, 2007.

[8] X. Wang and M. Lemmon, "On event design in event-triggered feedback systems," *Automatica*, vol. 47, no. 10, pp. 2319–2322, 2011.

[9] E. Garcia and P. J. Antsaklis, "Model-based event-triggered control for systems with quantization and time-varying network delays," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 422–434, 2012.

[10] A. Sahoo, H. Xu, and S. Jagannathan, "Near optimal event-triggered control of nonlinear discrete-time systems using neurodynamic programming," *IEEE transactions on neural networks and learning systems*, vol. 27, no. 9, pp. 1801–1815, 2015.

[11] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 2012.

[12] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE circuits and systems magazine*, vol. 9, no. 3, pp. 32–50, 2009.

[13] F.-Y. Wang, H. Zhang, and D. Liu, "Adaptive dynamic programming: An introduction," *IEEE computational intelligence magazine*, vol. 4, no. 2, pp. 39–47, 2009.

[14] G. A. Godfrey and W. B. Powell, "An adaptive dynamic programming algorithm for dynamic fleet management, i: Single period travel times," *Transportation Science*, vol. 36, no. 1, pp. 21–39, 2002.

[15] Y. Yang, K. G. Vamvoudakis, H. Modares, Y. Yin, and D. C. Wunsch, "Hamiltonian-driven hybrid adaptive dynamic programming," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.

[16] G. C. Goodwin and K. S. Sin, *Adaptive filtering prediction and control*. Courier Corporation, 2014.

[17] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, "Model-free q-learning designs for linear discrete-time zero-sum games with application to h-infinity control," *Automatica*, vol. 43, no. 3, pp. 473–481, 2007.

[18] K. G. Vamvoudakis, "Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems," *IEEE/CAA Journal of Automatica Sinica*, vol. 1, no. 3, pp. 282–293, 2014.

[19] A. Sahoo and S. Jagannathan, "Stochastic optimal regulation of nonlinear networked control systems by using event-driven adaptive dynamic programming," *IEEE transactions on cybernetics*, vol. 47, no. 2, pp. 425–438, 2016.

[20] V. Narayanan and S. Jagannathan, "Distributed adaptive optimal regulation of uncertain large-scale interconnected systems using hybrid q-learning approach," *IET Control Theory & Applications*, vol. 10, no. 12, pp. 1448–1457, 2016.

[21] A. Sahoo, V. Narayanan, and S. Jagannathan, "Optimal event-triggered control of uncertain linear networked control systems: A co-design approach," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2017, pp. 1–6.

[22] A. Sahoo, "Event sampled optimal adaptive regulation of linear and a class of nonlinear systems," 2015.

[23] D. Chatterjee and D. Liberzon, "Stability analysis of deterministic and stochastic switched systems via a comparison principle and multiple lyapunov functions," *SIAM Journal on Control and Optimization*, vol. 45, no. 1, pp. 174–206, 2006.

1 Conferences and Workshops

- IEEE International Conference on Cyber Physical and Social Computing (CPSCom 2020)
- IEEE Sensors Council Summer School 2020

2 Special Issues in Academic Journals

- IEEE Internet of Things Journal special issue on Security, Privacy, and Trustworthiness in Intelligent Cyber-Physical Systems and Internet-of-Things
- IEEE Transactions on Automation Science and Engineering special issue on Machine Learning for Resilient Industrial Cyber-Physical Systems
- SCIENCE CHINA Information Sciences special issue on Cyber-Physical Systems

Call for Contributions

Newsletter of Technical Committee on Cyber-Physical Systems (IEEE Systems Council)

The newsletter of Technical Committee on Cyber-Physical Systems (TC-CPS) aims to provide timely updates on technologies, educations and opportunities in the field of cyber-physical systems (CPS). The letter will be published twice a year: one issue in February and the other issue in October. We are soliciting contributions to the newsletter. Topics of interest include (but are not limited to):

- Embedded system design for CPS
- Real-time system design and scheduling for CPS
- Distributed computing and control for CPS
- Resilient and robust system design for CPS
- Security issues for CPS
- Formal methods for modeling and verification of CPS
- Emerging applications, e.g. automotive system, smart energy system, biomedical device, etc.

Please directly contact the editors and/or associate editors by email to submit your contributions.

Submission Deadline:

All contributions must be submitted by **July. 1st, 2021** in order to be included in the August issue of the newsletter.

Editor:

- Bei Yu, Chinese University of Hong Kong, Hong Kong byu@cse.cuhk.edu.hk

Associate Editors:

- Xianghui Cao, Southeast University, China xhcao@seu.edu.cn
- Long Chen, Sun Yat-Sen University, China chenl46@mail.sysu.edu.cn
- Caiwen Ding, University of Connecticut, USA caiwen.ding@uconn.edu
- Keke Huang, Central South University huangkeke@csu.edu.cn
- Yier Jin, University of Florida, USA yier.jin@ece.ufl.edu
- Subham Sahoo, Aalborg University, Denmark sssa@et.aau.dk

- Muhammad Shafique, Vienna University of Technology, Austria mshafique@ecs.tuwien.ac.at
- Umamaheswara Rao Tida, North Dakota State University, USA umamaheswara.tida@ndsu.edu
- Xiaolin Xu, Northeastern University, USA x.xu@northeastern.edu
- Ming-Chang Yang, Chinese University of Hong Kong, Hong Kong mcyang@cse.cuhk.edu.hk
- Junbo Zhao, Mississippi State University, USA junbo@ece.msstate.edu