

# Overview of the Colosseum: The World's Largest Test Bed for Radio Experiments

*David M. Coleman, Kenneth R. McKeever, Meredith L. Mohr, Lian R. Orraca Rosario, Kevin E. Parker, and Mark L. Plett*

## ABSTRACT

*As fifth-generation (5G) cellular technology emerges, it is apparent that the radio frequency (RF) spectrum is constrained by the ever-growing demands of application bandwidth and the number of devices vying for that bandwidth. Databases and procedures for managing the spectrum have become very complex and do not scale to meet today's on-demand spectrum requirements. In pursuit of novel methods to overcome these limitations, the Defense Advanced Research Projects Agency (DARPA) launched the Spectrum Collaboration Challenge (SC2) in 2016. The goal of the challenge, which culminated with the third and final competition in the fall of 2019, was to inspire participants to research, develop, and systematically test artificial intelligence algorithms across a network of radios to find the future paradigm for ensuring that the RF spectrum can support the bandwidths that next-generation applications will require. In support of SC2, the Johns Hopkins University Applied Physics Laboratory (APL) designed, developed, and hosted the Colosseum, the first-of-its-kind wireless research test bed in which competitors tested their algorithms and conducted their experiments in competition events. In addition to introducing SC2 and its goals, this article briefly describes the test bed architecture and the challenge events.*

## INTRODUCTION

In today's wireless world, an ever-increasing number of military and civilian devices contend for bandwidth in the radio frequency (RF) spectrum. "Managing this increasing demand while combating what appears to be a looming scarcity of RF spectrum is a serious problem for our nation."<sup>1</sup> To meet these challenges, the Defense Advanced Research Projects Agency (DARPA) introduced the Spectrum Collaboration Challenge (SC2) in 2016 to motivate research into new ways to ensure and expand access to the oversubscribed spectrum. The challenge was open to international participants (i.e.,

competitors) across academia and the commercial and defense industrial bases. In 2016, 30 teams spanning 5 countries signed up for the competition, and teams were down-selected through competition events in 2017 and 2018. Altogether, the competition awarded over \$17 million in prizes and fostered a new paradigm for collaborative spectrum access research.

To achieve challenge goals, competitors made use of advances in artificial intelligence (AI) and software-defined radios (SDRs). With the addition of AI in user equipment, negotiation of local spectrum allocations

occurs at machine speed. In this new paradigm, instantaneous environmental conditions and dynamic user demand drive spectrum allocation decisions, as opposed to the traditional model where licenses and policies set static frequency and bandwidth allocations. SC2 competitors have developed and tested novel strategies in pursuit of this collaborative intelligent radio network (CIRN) model “in which radio networks will autonomously collaborate and reason about how to share the RF spectrum, avoiding interference and jointly exploiting opportunities to achieve the most efficient use of the available spectrum.”<sup>1</sup>

To provide a controlled, realistic environment where competitors could test their experiments and compete in formal challenge events, APL designed, developed, and hosted the world’s largest wireless research test bed, referred to as the Colosseum. Located in a 30-foot by 20-foot server room (Figure 1) on APL’s Laurel, Maryland, campus until the final SC2 event, the remotely accessible platform was open to competitors over the internet 24 hours a day, 7 days a week. It allowed them to conduct large-scale experiments with intelligent radio systems in realistic scenarios and environments, such as a busy city plaza with conflicting wireless access points or an emergency that requires the rapid aggregation of multiple service providers.

This article introduces the architecture and capabilities of the Colosseum and discusses the official SC2 events, the first two of which were hosted at APL. Across the three events, DARPA awarded competitors over \$17 million in cash prizes to continue their research. After the third and final competition event in Los Angeles, California, in October 2019, the Colosseum was transitioned to Northeastern University as part of the National Science Foundation’s Platforms for Advanced Wireless Research (PAWR) program, where it will be installed in the university’s newly established Institute for the Wireless Internet of Things. The Colosseum will continue to be accessible for ongoing research in collaborative communication systems.



**Figure 1.** Photos of the Colosseum hosted in a server room on APL’s Laurel, Maryland, campus; it included 21 server racks comprising servers, SDRs, and network equipment. In October 2019 it moved to the Institute for the Wireless Internet of Things at Northeastern University.

## COLOSSEUM ARCHITECTURE (HOW IT WORKS)

The Colosseum wireless research test bed was designed to execute experiments autonomously and to provide three key end-user capabilities:

1. A testing environment where SC2 competitors could evaluate their designs using up to 128 radio nodes concurrently
2. A platform where competitors could informally scrimmage, allowing each the ability to practice within the test bed framework and exercise solutions with other competitors and incumbent radio systems
3. A playing field for formal test events at the end of each phase of the SC2 program

The test bed consisted of 128 independent standard radio nodes (SRNs). SRNs were each composed of a central processing unit (CPU), a graphical processing unit (GPU), and an SDR. The SDRs included a field-programmable gate array (FPGA) for local signal processing and were capable of exchanging Internet Protocol (IP) traffic among all radios “over the air,” via the RF Emulation System. Altogether, each SRN enabled competitors to develop and train AI algorithms for next-generation wireless networks in the absence of (or in addition to) a traditional spectrum access system.

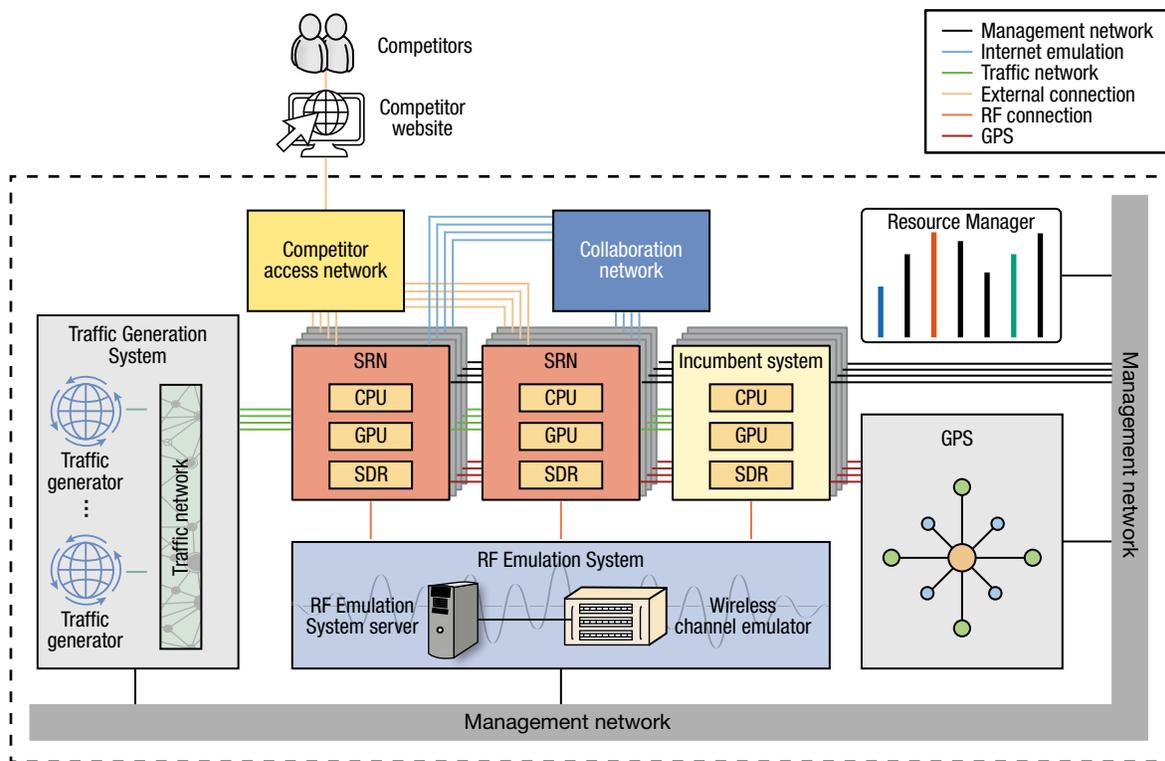
From the perspective of the competitors, the SRNs communicated with each other through the RF Emulation System. The team’s gateway node handled real-time spectrum allocation requests over a wired network connection, referred to as the collaboration network. This network represented fiber optic connections between cellular towers. Data exchanges between radios could occur only through RF communications on the connections established in the wireless channel emulator. Additionally, competitors were supplied a GPS daemon on each SRN so that they could estimate their positions in the virtual environment and the IP traffic flows they were required to satisfy throughout a scenario (see more on scenarios later in this article).

A complete architecture of the Colosseum is shown in Figure 2. In addition to the SRNs, there were five

key management entities: the competitor website, the Resource Manager, the GPS system, the RF Emulation System, and the Traffic Generation System. The website was the primary interface into the Colosseum and is where each competitor could request radio resources, select software to load on those resources, and view information on past, current, and future reservations. For each request of resources, the competitors accessed the SRNs via the competitor access network and requested Colosseum resources via the Resource Manager. For each Colosseum service request, the Resource Manager allocated resources and orchestrated the scenarios within the RF Emulation System, Traffic Generation System, GPS system, and incumbent systems. See the following subsections for brief descriptions of these subsystems, except the GPS service, which was a local process on the SRNs.

### Standard Radio Nodes

One of the major constructs of the CIRN framework is the SRN. Competitors used the Colosseum’s SRN subsystem to program the SDR, which was connected to a wireless channel emulator, giving them processing control in testing and executing their techniques for sharing and making the most efficient use of the RF spectrum. The SRN subsystem was capable of secure multiprocessing, enabling concurrent tests to run with-



**Figure 2.** Colosseum architecture. The test bed was made up of a collection of resources that competitors could use to design and test their solutions as well as resources that managed various functions within the test bed. All these resources worked together to enable competitors to test their solutions and DARPA to measure and assess them.

out interfering with each other. APL designed the SRN subsystem to operate under temporal constraints that minimized the deployment time while maintaining configurational and operational integrity between resource reservations. All data (logs, etc.) were available outside of the SRN containers so that competitors could post-process data after running experiments and DARPA could use the data to measure team's success. See the article by White et al. in this issue for more on SRNs.

### Resource Manager

The Colosseum's custom Resource Manager satisfied three main requirements. The first was to ensure fairness across competitors by allocating resources through a token system and scheduling. In this way, competitors had equal opportunity to execute experiments on a weekly basis. The second was to conduct automated experimentation and orchestration through a process known as automated mode. In automated mode, the Resource Manager synchronized the individual services used within an experiment (e.g., RF emulation, traffic generation, GPS, collaboration network; these services are described in more detail throughout this issue). Finally, the Resource Manager provided a central point for verification of Colosseum operations by maintaining a detailed log of operations. DARPA was able to use this log to audit Colosseum operations so that it could verify competition data. This subsystem is described in more detail in the article by Mok et al. in this issue.

### Incumbents

Another critical component of the Colosseum was its legacy radio emulators, referred to as Colosseum incumbents. These incumbents emulate the RF behavior of existing real-world radio systems already in place, such as radars for weather monitoring, receivers for spectrum monitoring, and battlefield jammers. Because these incumbent systems will remain in use for a long time, intelligent radios of the future will need to operate around and adapt to them to avoid interfering with them. SC2 scenarios featured three types of incumbent systems: passive receivers, radars, and noise jammers. These incumbent systems served as RF obstacles; SC2 competitors had to find innovative approaches to detect and work around them while simultaneously administering their own communications for maximum data throughput efficiency. See the article by Yim et al. in this issue for more on incumbents.

### Traffic Generation System

The Traffic Generation System emulated realistic traffic flows between radios to facilitate measurement and evaluation of competitors' designs for sharing the RF spectrum. The subsystem enabled on-demand generation, logging, and analysis of Internet Protocol (IP) ver-

sion 4 (IPv4) traffic in the Colosseum. IP traffic provides good metrics for the level of success of a competitor radio because IP packets can be counted, and statistics such as bit rate throughput, latency, jitter, and loss can be calculated. These metrics helped competitors improve the performance of their algorithms after practice runs, ensured connectivity of the nodes, and provided an application layer that the nodes had to service. In addition, DARPA used the metrics during competitions to evaluate and compare the performance of competitors' radios. See the article by Curtis et al. in this issue for more on the Traffic Generation System.

### RF Emulation System

Barcklow et al., in an article in this issue, describe the Colosseum's RF Emulation System. This subsystem mimicked real-world phenomenon such as propagation delay, Doppler shift, and power attenuation among the full complement of the Colosseum's antennae (128 two-channel radios, or 65,536 wireless communications channels). Made up of a server that read and processed scenario files and a wireless channel emulator, the RF Emulation System simulated isolated virtual environments across multiple concurrent experiments, enabling challenge competitors to research and develop their next-generation AI solutions for wireless network systems.

All these subsystems worked together to enable SC2 competitors to design, test, and execute their solutions in the context of predefined scenarios, described in more detail below.

## SC2 SCENARIOS

SC2 centered on a series of scenarios designed to mimic real-world challenges a network of collaborative autonomous radios would have to overcome. "These custom RF scenarios consist[ed] of 3-D models of the environment and the motion of all the radios. From this, a toolchain automatically generate[d] the terabytes of data that describe the changing characteristics of radio wave propagation between each pair of radios as they move. This data [was] streamed into Colosseum in real-time to drive the experiment"<sup>2</sup> in scrimmages, practice runs, and each formal event. Competitors used these scenarios to validate their algorithms, while DARPA used them to evaluate overall performance in the competition. Scenarios included common wireless communication challenges in realistic environments, such as

1. operating in close proximity and with limited bandwidth (e.g., in a residential neighborhood);
2. adapting to temporal surges in spectrum demand (e.g., from wireless hot spots);
3. mitigating interference to a legacy radio system (e.g., a radio broadcast station);



the Colosseum's modes of operation in more detail in the article on the Resource Manager, in this issue.

## COMPETITION EVENTS

The challenge consisted of several events: scrimmages, two formal preliminary events, and the final formal event. These events allowed competitors to develop, test, and refine their techniques, and the formal events enabled DARPA to assess competitors' performance.

### Scrimmages

Conducted by APL before each formal event, scrimmages gave competitors the opportunity to build their radios up to the specifications required for participation in the formal events. Scrimmages generally ran every 4–6 weeks on dates that had been determined at the beginning of each year of the challenge. Before participating in scrimmages, competitors received practice versions of the scenarios so that they could test their algorithms against them. All scrimmages included scenarios intended to be used in the upcoming formal event. The scrimmage process was similar to that for a formal event: in the first round, APL collected and validated competitor radio and algorithm submissions. The competitors with valid radio and algorithm submissions were paired with other valid competitors. The matches were executed and the resulting logs were then distributed to the teams so that they could run their own analyses on their scrimmage data.

### Preliminary Event 1

APL hosted the first preliminary event on December 12, 2017. To qualify for the event, teams had to prove (validate) that their radio designs could operate in today's wireless paradigm—in other words, that their radio networks could transmit data in an environment free of interference. After having spent months developing their designs and refining them after competing in scrimmages, 19 teams qualified for and participated in the event. The Colosseum ran 475 autonomous matches comprising six scenarios that simulated the competition for bandwidth seen in today's actual electromagnetic environments. "The competing teams faced fluctuating bandwidths and interference from other competitors as well as DARPA designed bots that tested and challenged their radio designs."<sup>3</sup> Every team had 5 radios, for a maximum total of 15 team radios per match. Of the 19 teams, 10 teams were judged to have best collaborated to share the spectrum and were awarded \$750,000 each in prize money.

### Preliminary Event 2

The second preliminary event, hosted at APL a year later on December 12, 2018, "further challenge[d] competitors with an interference environment beyond what

existing commercial and military radios can handle—upping the number of simultaneous wireless network types from three to five, and raising the total number of radios from 15 to 50."<sup>3</sup> The event included 15 teams, spanned 3 rounds and over 450 competition runs, and ended with DARPA awarding \$6 million dollars in prize money. Leveraging automation in the Colosseum, the competition runs were executed in parallel over a 4-day period leading up to the hosted event. In the first round, competitor radios were tested against civilian applications in a shared spectrum. Challenges in these scenarios included multiple Wi-Fi hot spots in a retail plaza, intelligent spectrum algorithms in the presence of legacy (i.e., incumbent) systems, and a citywide military operation with progressively complex communications requirements and obstacles.

From these scenarios, the top 12 teams advanced to the second round. In this round, competitors were tested against scenarios that included disaster relief efforts and battlefield jamming.<sup>3</sup> The number of teams within a scenario also increased from three to five, while the available spectrum remained identical, challenging each radio to share a heavily constrained resource. From these scenarios, the top eight teams advanced to the final round. In the final round, each of the teams had to demonstrate the utility of their intelligent radios. Each radio was re-subjected to a previous scenario but was required to operate in both intelligent and legacy modes of operation. To demonstrate the effectiveness of their intelligence algorithms, a team was required to achieve a higher score when using the intelligence than when using the legacy (non-intelligent) mode of operation, proving that their work had improved spectrum use and removed the need for rigid spectrum management in future systems.

### Final Event

All teams that competed in the second preliminary event qualified for the final event. Before the event, teams competed in a championship play-in round to decide the 10 teams that would compete in the final event.<sup>4</sup> The final event was held in front of a live audience at the 2019 Mobile World Congress Los Angeles on October 23, 2019. The Colosseum was transported from APL to California for the culmination of the 3-year competition so that the 10 qualifying teams could compete in real time in Los Angeles during the last round.

The structure of the finale was similar to that of preliminary event 2 but with a few surprises. In six rounds of competitive play, teams' AI-enabled radios were subjected to five different scenarios inspired by real-world situations: a military mission where soldiers supported communications while sweeping an urban area; a response to a natural disaster that forced teams to act as disaster relief agencies and work with the available ad hoc communications; a wildfire response effort that involved

prioritizing communications with aerial firefighters over other needs; a high-traffic shopping area where teams had to manage surges in traffic at various stores as their radios acted as Wi-Fi hot spots; and the “trash compactor,” which reduced the amount of available spectrum, narrowing it from 20 MHz to 10 MHz to 5 MHz.<sup>5</sup> The scenarios were designed to test different characteristics of competitors’ systems, such as their ability to provide stable service, prioritize various types of wireless traffic, and cope with extremely congested environments.

The final five teams advanced to the sixth and final round, facing modified versions of the five previous scenarios that incorporated new obstacles to overcome, such as incumbent radio systems, which are known to be sensitive to interference. With the Colosseum on-site, the results were able to be presented in real time. The winning team, University of Florida’s GatorWings, was awarded the \$2 million grand prize. MarmotE finished second and received \$1 million, and the third-place prize of \$750,000 went to Zylinium.

## PROJECT MANAGEMENT AND DEVELOPMENT OPERATIONS

The design and management of the Colosseum required rigorous planning and monitoring as well as close coordination among APL team members and between APL and DARPA. As the lead in development, integration, and operation of the Colosseum, APL had to adhere to a tight development schedule and then, after the Colosseum was released, ensure continuous access for competitors while also incorporating fixes and enhancements to the test bed. To meet these goals, APL applied tested project management tools and techniques and an agile framework. Some of the tools APL used to enhance productivity included Jira for tracking and collaborating, Confluence for housing documentation, Slack for facilitating quick and organized communication, and Jenkins for continuous integration and delivery of software. The article by Freeman et al. in this issue describes the project management approach and tools in more detail.

Complementing the project management principles, a development and operations (DevOps) approach guided the teams developing software codebases, deploying system configurations, and monitoring the status of hardware systems. DevOps represents a change in information technology (IT) culture, focusing on rapid IT service delivery through the adoption of agile, lean practices in the context of a systems-oriented approach. DevOps emphasizes people (and culture) and seeks to improve collaboration between operations and development teams. DevOps implementations use technology, especially automation tools that can leverage an increasingly programmable and dynamic infrastructure from a life cycle perspective. DevOps is fundamentally

the merging of two disciplines—software development and system administration. The article by Plummer and Taylor details the DevOps process and tools the team used to build and maintain the Colosseum.

## CONCLUSION

DARPA’s SC2 motivated competitors to research, develop, and test AI solutions to usher in the future of RF spectrum management. At the heart of SC2 was the APL-designed, -developed, and -hosted Colosseum, the world’s largest wireless research test bed for radio experiments. Using the Colosseum’s robust set of integrated resources, SC2 competitors were able to continuously develop and test their designs in a repeatable environment. Using the vast data the Colosseum provided, DARPA was able to measure how well the designs performed during formal competition events. APL relied on various project management tools and techniques and a DevOps approach to ensure the success of the Colosseum and, ultimately, the SC2 program. After the third and final SC2 competition in October 2019, the Colosseum transitioned to Northeastern University, where it will have an enduring impact on spectrum research and collaboration, serving future researchers in collaborative communication systems.

**ACKNOWLEDGMENTS:** We thank Paul Tilghman (DARPA SC2 program manager) and Craig Pomeroy and Kevin Barone (Systems Engineering and Technical Assistance at DARPA) for their invaluable collaboration and support. We also thank the many APL SC2 contributors, whose names are listed on the inside back cover of this issue of the *Digest*. This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA). The views, opinions, and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the US government.

## REFERENCES

- <sup>1</sup>P. Tilghman. “Spectrum Collaboration Challenge (SC2).” DARPA SC2 website. <https://www.darpa.mil/program/spectrum-collaboration-challenge> (accessed Jul. 18, 2019).
- <sup>2</sup>“The technology.” DARPA SC2 website. <https://www.spectrumcollaborationchallenge.com/about/technology/> (accessed Jul. 18, 2019).
- <sup>3</sup>“Preliminary event #1 recap.” DARPA SC2 website. <https://www.spectrumcollaborationchallenge.com/preliminary-event/> (accessed Jul. 18, 2019).
- <sup>4</sup>“DARPA announces final teams to compete in Spectrum Collaboration Challenge Championship Event.” DARPA SC2 website. <https://www.spectrumcollaborationchallenge.com/darpa-announces-final-teams-to-compete-in-spectrum-collaboration-challenge-championship-event/> (Sep. 10, 2019).
- <sup>5</sup>“DARPA grand challenge finale reveals which AI-managed radio system shares spectrum best.” DARPA SC2 website. <https://www.spectrumcollaborationchallenge.com/darpa-grand-challenge-finale-reveals-which-ai-managed-radio-system-shares-spectrum-best/> (accessed Nov. 20, 2019).



**David M. Coleman**, Asymmetric Operations Sector, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

David M. Coleman is research area lead for spectrum operations and supervisor of the Communication Systems Concepts Section in APL's Asymmetric Operations Sector. He earned a BS in computer engineering from Elizabethtown College, an MS in computer engineering from the University of Arkansas, and a PhD in electrical engineering from University of Maryland, College Park. For the DARPA SC2 effort, Dr. Coleman led the software development for the Colosseum's RF Emulation System as well as the Interface Control Document definition between this subsystem and the Colosseum. Dr. Coleman has presented at many conferences and has been published in journals and proceedings. His email address is david.coleman@jhuapl.edu.

**Kenneth R. McKeever**, Asymmetric Operations Sector, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Kenneth R. McKeever is the acting assistant group supervisor of the Wireless Cyber Capabilities Group in APL's Asymmetric Operations Sector. He has a BS in electrical engineering from Penn State and an MS in electrical and computer engineering from Johns Hopkins University. His background and experience is in wireless security, communication systems theory, military and commercial communication technologies and protocols, signal processing, and computer networking. He has contributed to design, simulation, and both laboratory and field test and evaluation of various systems. As acting assistant group supervisor, Ken assists with directing and developing the group, including developing staff, prioritizing tasking, reaching out to and collaborating with other groups, and identifying and fostering new, impactful opportunities. He is a member of IEEE. His email address is kenneth.mckeever@jhuapl.edu.



**Meredith L. Mohr**, Asymmetric Operations Sector, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Meredith L. Mohr is a member of the Senior Professional Staff and supervisor of the Wireless Cyber Access Section in APL's Asymmetric Operations Sector. She holds a BS in electrical and computer engineering from North Carolina State University and an MS in electrical and computer engineering from Johns Hopkins University. Her technical background includes signals and systems, satellite communication, programming, and data processing. Meredith has experience with RF test and evaluation, communication systems, and systems engineering. She is a member of IEEE. Her email address is meredith.mohr@jhuapl.edu.



**Lian R. Orraca Rosario**, Asymmetric Operations Sector, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Lian R. Orraca Rosario is an electrical engineer in APL's Asymmetric Operations Sector. He earned a BS in electrical engineering from the University of Puerto Rico at Mayaguez and is working toward an MS in electrical and computer engineering from Johns Hopkins University. Lian has experience in Python scripting, C and C++ programming, Java, MATLAB, Bash, hardware design, RF communication systems, and RF engineering. He was part of the event team for DARPA SC2. This team was in charge of validating all the competitor radios, running them against each other in different scenarios in Colosseum, and sending information to the scoring engine while protecting the software and logs of each team from the others. Lian was also the principal investigator for an independent research and development project focused on applying machine learning techniques to generate a local model of RF transmitter behavior and predict future transmission, and presented a paper about compact high-frequency receiver system design at the IEEE Radio and Wireless Symposium. His email address is lian.orraca@jhuapl.edu.



**Kevin E. Parker**, Asymmetric Operations Sector, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Kevin E. Parker is a project manager in APL's Asymmetric Operations Sector. He holds a BSEE from Virginia Tech and an MSEE from Johns Hopkins University. Kevin started his career by serving on submarines in the US Navy. After 7 years of active duty, he came to APL while continuing in the Navy Reserve. Over his 31-year career at the Lab, Mr. Parker's technical work has focused primarily on communication systems engineering and receiver signal processing. More recently he has served in a variety of technical, project, and line leadership roles to help advance support to the cyber operational forces. Kevin served as the project manager for the DARPA SC2 project. His email address is kevin.e.parker@jhuapl.edu.



**Mark L. Plett**, Asymmetric Operations Sector, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Mark L. Plett is supervisor of APL's Wireless Cyber Capabilities Group from 2013 until 2019. He earned BS degrees in electrical engineering and physics, an MS in electrical engineering, and a PhD in electrical engineering, all from the University of Maryland. He is an expert in communications and radio frequency wireless systems. He joined APL in 2009, and throughout his time at the Lab, he made critical impacts on the development of wireless systems leveraging advanced technologies such as software-defined radio. His technical breakthroughs in wireless systems enabled several sponsor-critical cyber operations missions. Before joining APL, Mark worked as a senior hardware developer for Microsoft and as a senior systems analyst for Terabeam.