58th IAA SYMPOSIUM ON SAFETY, QUALITY AND KNOWLEDGE MANAGEMENT IN SPACE ACTIVITIES (D5)
Cybersecurity in space systems, risks and countermeasures (4)

Author: Dr. Tim Parsons
Australia

Mr. Bruce Tulloch
IEEE member, United States

DEEPFAKES AND GENERATIVE AI: AN OLD THREAT TO THE SPACE ECONOMY IS NEW AGAIN

Abstract

In a 2019 survey, only 61 percent of US respondents strongly believed that the 1969 moon landings really happened. More recently, researchers have shown how an off-the-shelf computer system worth less than $1,000 can generate photorealistic, credible and fake high-resolution video and audio of a person which neither a human nor an AI can distinguish from real video or audio of that person. We now live in an era where thanks to Generative AI, reality itself is under threat.

In this paper, we discuss ways in which data from Space, specifically images, data and video generated by Earth Observation, Positioning and Timing, and Deep Space Exploration satellites and systems missions, can and will be convincingly faked by bad actors using Generative AI. The consequences of this are likely to be catastrophic, and with critical decisions already being made by governments, corporations and individuals on Earth based on data from increasingly commercial space systems. At best, the entire value of data from space will be called into question.

To stimulate debate on mitigating this threat, we present case studies of initiatives from multiple sectors that counter generative AI abuse. These include standards-setting bodies like C2PA (Coalition for Content Provenance and Authenticity) and TEIA (Trusted Energy Interoperability Alliance), which define technical and governance standards for verifiable data provenance across global value chains. We also examine technical infrastructure like the Hedera Hashgraph public ledger, now under the Linux Foundation, which verifies data integrity between source and viewer. Finally, we explore optical security utilizing quantum technologies (QKD, BB84 algorithms) and cryptographic schemes that protect high-bandwidth optical communications from interception or data injection.

This paper will also provide links to open-source Github repositories where some of these emerging standards are being expressed in deployable code, example datasets, and implementation patterns, on devices ranging embedded sensors at the edge to high-performance compute infrastructure in the cloud.

The authors hope these detailed case studies will stimulate major players and interested parties to come together now to develop protocols for provenance and authenticity to underpin the integrity of data from space, so that there is a community-wide answer to the first incidents of data purporting to be from space, but actually generated by AI, causing real harm on Earth.