SPACOMM 2019

The Eleventh International Conference on Advances in Satellite and Space Communications

ISBN: 978-1-61208-694-1

March 24 - 28, 2019

Valencia, Spain

SPACOMM 2019 Editors

Timothy Pham, Jet Propulsion Laboratory, USA

SPACOMM 2019

Forward

The Eleventh International Conference on Advances in Satellite and Space Communications (SPACOMM 2019), held between March 24, 2019 and March 28, 2019 in Valencia, Spain, continued a series of events attempting to evaluate the state of the art on academia and industry on the satellite, radar, and antennas based communications bringing together scientists and practitioners with challenging issues, achievements, and lessons learnt.

Significant efforts have been allotted to design and deploy global navigation satellite communications systems, Satellite navigation technologies, applications, and services experience still challenges related to signal processing, security, performance, and accuracy. Theories and practices on system-in-package RF design techniques, filters, passive circuits, microwaves, frequency handling, radars, antennas, and radio communications and radio waves propagation have been implemented. Services based on their use are now available, especially those for global positioning and navigation. For example, it is critical to identify the location of targets or the direction of arrival of any signal for civilians or on-purpose applications; smarts antennas and advanced active filters are playing a crucial role. Also progress has been made for transmission strategies; multiantenna systems can be used to increase the transmission speed without need for more bandwidth or power. Special techniques and strategies have been developed and implemented in electronic warfare target location systems.

We welcomed academic, research and industry contributions. The conference had the following tracks:

- Satellite and space communications
- Satellites and nano-satellites
- Satellite/space communications-based applications

We take here the opportunity to warmly thank all the members of the SPACOMM 2019 technical program committee, as well as all the reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to SPACOMM 2019. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

We also thank the members of the SPACOMM 2019 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that SPACOMM 2019 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the domain of satellites and space communications. We also hope that Valencia, Spain provided a pleasant environment during the conference and everyone saved some time to enjoy the historic charm of the city.

SPACOMM 2019 Chairs

SPACOMM 2019
Committee

SPACOMM Steering Committee
Timothy T. Pham, Jet Propulsion Laboratory - California Institute of Technology, USA
Stelios Papaharalabos, LN2 dB, Greece
Cathryn Peoples, Ulster University & The Open University, UK

SPACOMM Industry/Research Advisory Committee
Vittorio Dainelli, Rheinmetall Italia S.p.A., Italy
Brian Niehoefer, TÜV Informationstechnik GmbH, Germany

SPACOMM 2019 Technical Program Committee

Ashfaq Ahmed, COMSATS Institute of Information Technology - Wah campus, Pakistan
Mohamed Al-Mosawi, University of Portsmouth, UK
David N. Amanor, INTEL CORPORATION, Hillsboro, USA
Şuayb Ş. Arslan, MEF University, Istanbul, Turkey
Michael Atighetchi, Raytheon BBN Technologies, USA
Arash Behboodi, RWTH Aachen University, Germany
Mark Bentum, University of Twente, Netherland
Maria Dolores Cano Baños, Universidad Politécnica de Cartagena. Spain
Genshe Chen, Intelligent Fusion Technology Inc., USA
Yao-Yi Chiang, University of Southern California, USA
Vittorio Dainelli, Rheinmetall Italia S.p.A., Italy
Raed Daraghma, Palestine Technical University (PTU), Tulkarm, Palestine
Peng Deng, The Pennsylvania State University, USA
Alisa Devlic, Huawei Swedish Research Center, Sweden
Suleyman Eken, Kocaeli University, Turkey
Zabih Ghassemlooy, Northumbria University, UK
Gregory Hellbourg, University of California, Berkeley, USA
Ibrahim Hokelek, TUBITAK BILGEM, Turkey
Huiping Huang, Southern University of Science and Technology of China, Shenzhen, China
Suk-Seung Hwang, Chosun University, Korea
Adil Hakeem Khan, Nation College of Engineering and Tech., Guna MP, India
M. Michael Kobayashi, NASA Jet Propulsion Laboratory, USA
Arash Komaee, Southern Illinois University, USA
Emmanouel T. Michailidis, University of Piraeus, Greece
Kresimir Malaric, University of Zagreb, Croatia
Luca.Martino, Universitat de Valencia, Spain
Lyudmila Mihaylova, University of Sheffield, UK
Ralf Mueller, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany
Brian Niehoefer, TUV Informationstechnik GmbH, Germany

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

System Development and Spacecraft Testing of the Morehead State University Ground Station

Timothy Pham, Jason Liao. E.J.Wyatt
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California, USA
e-mail: {Tim.Pham, Jason.Liao, E.J.Wyatt}@jpl.nasa.gov

Benjamin Malphrus, Jeffrey Kruth, Robert Kroll, Michael.Combs, Tobias Gedenk, Sarah Wilczewski, Jacob Lewis, Alex Roberts, Emily Mattle, Chloe Hart
Space Science Center, Morehead State University
Morehead, Kentucky, USA
e-mail: {b.malphrus, j.kruth, r.kroll, m.combs, tgedenk, sewilczewski, fjlewis, ajroberts3, ejmattle, cmhart}@moreheadstate.edu

Abstract— **In support of the Lunar IceCube mission, as well as other CubeSats to be launched on the upcoming NASA Exploration Mission 1, the 21-m ground antenna at the Morehead State University in Kentucky, United States is under development to augment its telemetry, tracking and command capability at X-band (previous capability is limited to S-band). The new system is expected to be operational in 2020. This is the first collaboration between National Aeronautics and Space Administration's Deep Space Network (DSN) and university in ground system development. It will result in a university's tracking station fully compatible with the DSN and thus increase DSN support to mission users with more antenna assets. The new hybrid system architecture combines commercially available RF components, custom-designed microwave feed for Morehead antenna, and specialized DSN digital equipment. Such an approach minimizes the implementation cost, offers learning opportunity for the students to gain experience in system design and implementation, and maximizes system capability and compatibility with the DSN. The ground station at the Morehead State University will handle signal transmission to and reception from spacecraft, and other signal processing functions such as modulation/demodulation, decoding, etc. The backend equipment at the Jet Propulsion Laboratory in Pasadena, California will be responsible for data interface to the mission users. This architecture makes the Morehead antenna appears as one of the DSN antennas. Since early 2018, the 21-m ground system at Morehead has been upgraded with capability to receive spacecraft signal at X-band and process its telemetry and Doppler data. In mid 2019, the Morehead system will be further enhanced with a transmit capability for spacecraft command and ranging. Testing with spacecraft currently in operation such as Osiris Rex, MAVEN and Hayabusa2 is ongoing in an effort to validate system capability and characterize system performance. Preliminary results indicate that the downlink system is functional. They also enable assessment of the Gain-to-System Noise Temperature ratio (G/T) of the Morehead 21-m antenna.**

Keywords - Morehead State antenna; ground system; Lunar IceCube.

I. INTRODUCTION

CubeSats and other small satellites are increasingly used for Earth remote sensing, for science research, and for unique communications activities, all of which have increasing data throughput requirements. In addition, CubeSats are being planned for interplanetary research, with 13 CubeSats slated to fly on the NASA's Exploration Mission-1 (EM-1) in 2020, opening the door for CubeSat and smallsat exploration of the solar system. As these CubeSat missions venture to the distance of the moon and beyond to other bodies within the Solar System, they require a ground tracking system with greater capabilities than previously required for low Earth orbiters. Performance attributes in the ground system such as large antenna, high gain efficiency, operation at higher performing X-band frequency, low noise and low-loss equipment, most efficient forward error correction coding, high transmitting power, etc. become critical to the communications with deep space spacecraft.

Given the expected significant increase in CubeSat missions, beyond what is currently supported by the NASA Deep Space Network (DSN), the NASA Advanced Exploration Systems (AES) Program has been funding an implementation at the Morehead State University to enable its 21-m antenna to support the Lunar IceCube mission, as well as other EM-1 CubeSats. Leveraging on the expertise in deep space communications of the Deep Space Network, a partnership between the Jet Propulsion Laboratory (JPL) in California, United States and the Morehead State University (MSU) in Kentucky, U.S., was established to aid the development of ground station at Morehead, and to develop a strategy that would enhance DSN capabilities by utilizing existing non-NASA assets (i.e., university and non-profit radio astronomy observatories). Our goal is to build a ground station that is capable of deep space communications and tracking, with maximum compatibility with the DSN, and within limited budget. The approach we took is to optimally combine specialized DSN equipment with those available commercially. This hybrid system minimizes

 ISBN: 978-1-61208-694-1

development cost; saving the non-recurring engineering cost on signal processing equipment already exist in the DSN. The rest of the system is designed by the Morehead State technical staff and students, from commercially available products; thus, offers a learning opportunity to the students and, at the same time, meeting the low-cost objective.

In Section 2 of this paper, we describe the Morehead ground system architecture that strives for maximum capability with a minimum cost, using the hybrid approach mentioned above. Such an architecture allows the MSU ground system to serve as a node on the Deep Space Network, which could help to offset the tracking load on the DSN for certain class of missions such as CubeSats. In Section 3, we present a nominal operational concept on how the system is expected to operate. Telemetry, tracking and command data flows will be described, along with service management aspects such as the antenna scheduling and generation of predicted Doppler frequencies, antenna pointing, and signal conditions. We highlight key system performance metrics in Section 4. Finally, Section 5 captures some preliminary test results with a few spacecraft currently in operation, such as Osiris Rex, MAVEN and Hayabusa2.

II. SYSTEM ARCHITECTURE

Prior to this implementation effort, which started in 2016, the Morehead State University 21-m antenna has been used to support several educational research picosatellites and nanosatellites such as KySat-1, KySat-2, and the Cosmic X-band Background Nanosat series (CXBN and CXBN-2). The system was also used for short-term capability demonstrations with NASA flight mission such as the Lunar Reconnaissance Orbiter and the Thermosphere, Ionosphere, Mesosphere Energetics and Dynamics (TIMED) [1]. Most recently, the system served as the primary ground station for the Arcsecond Space Telescope Enabling Research in Astrophysics (ASTERIA) mission [2], funded by the Jet Propulsion Laboratory (JPL) in collaboration with the Massachusetts Institute of Technology (MIT). The 21-m antenna, prior to this upgrade, was equipped with a replaceable feed that can support operations at various frequencies in the Radio Frequency (RF), specifically the UHF-band (400- 470 MHz), L-band (1.4 – 1.7 GHz), S-band (2.2 – 2.5 GHz), C-band (4.8 - 5 GHz), X-band (7.0 – 7.8 GHz), and Ku-band (11.2 – 12.7 GHz). Signal reception is available at all these frequency bands and transmission is available at UHF and S-band.

For the EM-1 CubeSat missions, communications between flight and ground systems will be done at X-band to maximize the link performance. The X-band capability will be supported by new equipment (antenna feed, transmitter, receiver, network connections, frequency reference, etc.) and new operational processes (antenna scheduling, spacecraft pointing and doppler prediction, data delivery, etc.) described in this paper. The only thing that is common with the previous S-band and UHF systems is the common infrastructure such as antenna and electrical power. The ground system will have new equipment to transmit signal at 7.1-7.2 GHz and to receive spacecraft signal at 8.4-8.5 GHz. A new X-band antenna feed will need to be developed. On the uplink, a 2-kW transmitter will be deployed, along with the exciter electronics for generation of command and ranging signals. On the downlink, a cryogenic low noise amplifier operating at a very low noise temperature of 11 Kelvins will help to maximize the signal detection. A low-loss DSN-based receiver with its associated telemetry decoding and ranging processors that are optimized for low Signal-to-Noise Ratio (SNR) condition normally seen in deep space communications will also be deployed. These enhancements, relative to capability in previous missions, extend the communications to spacecraft that travel to the Moon and beyond – a 1000 times further than the low Earth orbits.

For past mission support, the MSU ground station could transmit telecommand to spacecraft and receive telemetry data; however, there was no ranging measurement required for the navigation of these nanosatellites. For future missions like the Lunar IceCube, ranging capability is essential in navigating the spacecraft to the moon. Another new challenge is the use of highly efficient error correcting codes, such as Turbo codes, which perform very close to the Shannon limit of the information channel capacity. Unfortunately, there was difficulty in finding commercial products that could support these two functions within available budget. As a result, a hybrid architecture that merges the university-developed hardware with some DSN-developed signal processing components was selected. The university focuses on RF analog components, using commercial parts available from other projects, as well as some newly procured components. The digital portion of the system that process command, telemetry and radiometric (Doppler and ranging) are replica of DSN equipment. This architecture enables the MSU ground station to act as a DSN node of operations since it has the same data interfaces to mission users as other DSN antennas. Since the DSN-provided equipment supports the data interfaces in compliance with the Consultative Committee for Space Data Systems (CCSDS) specifications, the MSU system has an inherent benefit of interoperability with other ground stations worldwide that are CCSDS compliant.

Figure 1 shows the new architecture of the Morehead ground station that is being developed. The Mission Operation Center (MOC) interfaces with the Morehead Station ground station via the DSN Deep Space Operation Center (DSOC). Command data could be sent from the MOC directly to the Uplink (UPL) equipment at MSU. Telemetry processing at MSU produces the received telemetry frames, which are then relayed to the Telemetry Tracking Delivery (TTD) at JPL before being delivered to the MOC. The same delivery occurs with radiometric data of Doppler and ranging measurements. The network connection between Morehead State University and the Deep Space Operation Center at JPL is via the NASA

 ISBN: 978-1-61208-694-1

Mission Backbone Network. This network connection has high redundancy and reliability, making it best suited for mission operations. The selected bandwidth of this connection is based on considerations of expected data rate needed for mission support and the annual rental cost. The Lunar IceCube and other EM-1 CubeSats have a maximum data rate under 1 Mbps. A decision was made to set the leased bandwidth at 5 Mbps to provide some flexibility for expanded operational needs.

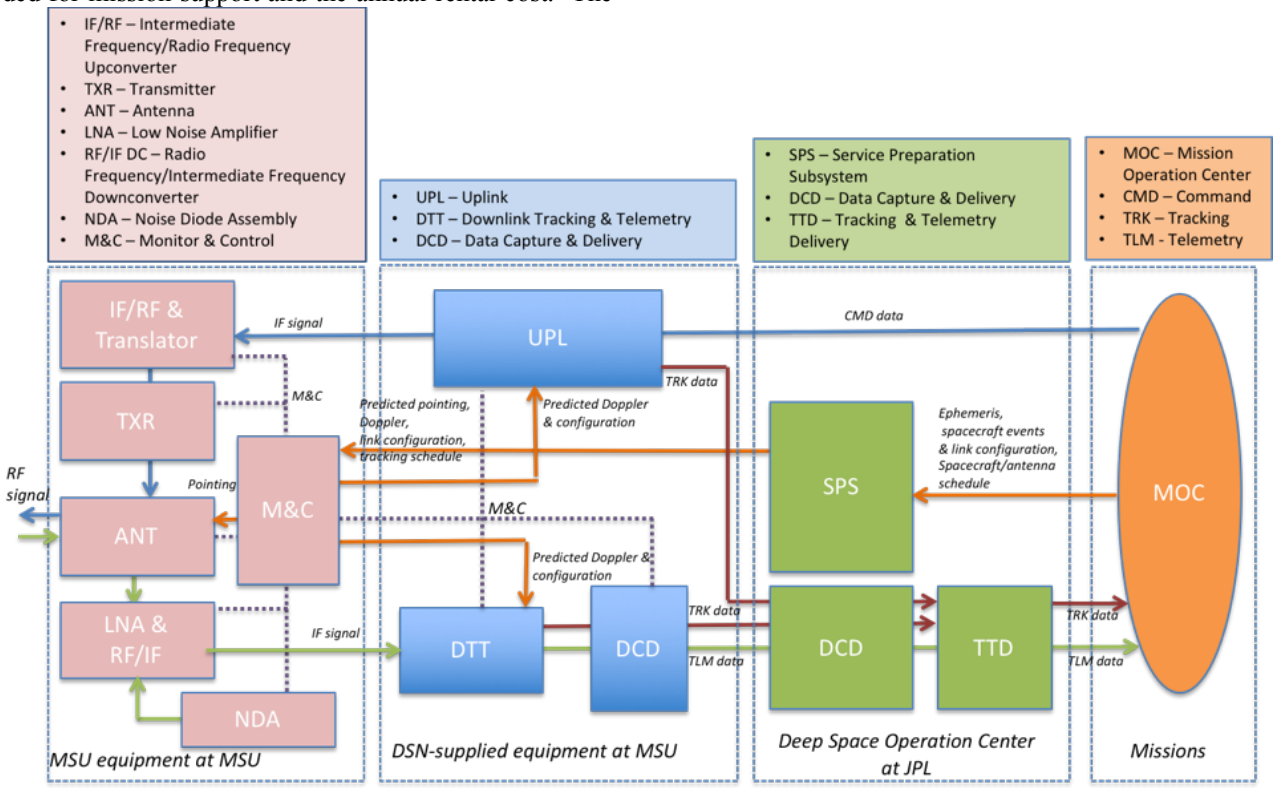


Figure 1. System architecture for the new X-band capability

There is also other ancillary equipment needed for ground station testing, calibration and operation monitoring, besides those used for telecommand and telemetry. The Test Translator with a standard CCSDS transponder ratio of 880/749 is used to convert the uplink signal from a transmitted frequency of ~7.2 GHz to a received signal at ~8.4 GHz. With the Translator loopback, the ranging delay within the ground system can be precisely measured. This station delay is then removed from the ranging measurements observed during spacecraft tracking in order to properly determine the spacecraft range. The loop back capability also enables a partial verification of the uplink and downlink equipment. By configuring the uplink equipment to generate a command data stream, feeding that signal to the downlink equipment, and being able to extract telemetry data, proper operation of uplink signal generation and downlink telemetry processing can be verified.

Another component that also aids with the monitoring of system performance is the Noise Diode Assembly (NDA). This equipment injects a known noise power into the system at a periodic frequency. By looking at the difference in power levels with and without the added noise, one can determine the noise temperature of the ground system. At MSU, such calibration is expected to be undertaken at the start of the pass, rather than continually through the pass as it is done in the DSN antennas. This is a trade-off for a simpler design in the noise calibration.

III. Operational Concept

In this section, we first describe the data interfaces with mission users and their spacecraft. Then, we'll focus on service management aspects of the operations, which include preparation for the tracking pass, as well as equipment monitoring during the pass.

For telemetry processing, as indicated in Figure 1, the signal received at the 21-m antenna is amplified by the cryogenic low noise amplifier and down-converted to an IF frequency around 300 MHz. The signal is then routed from the antenna to the mission control room, about 1 km away, via the fiberoptic link. The IF signal is then fed into an FPGA-based receiver that will digitize, demodulate and decode the signal. The extracted telemetry frames are sent to JPL by the Data Capture and Delivery (DCD) assembly. As the name implies, the DCD also captures and archives the data in short term, giving the option to retransmit the data should the network connection to JPL be temporarily down. At JPL, the telemetry frames are delivered to the mission operation system via an interface compliant with the CCSDS

 ISBN: 978-1-61208-694-1

Space Link Extension (SLE) of Return All Frames or Return Channel Frames. Telemetry frames can also be further processed at JPL into packets or file products under the CCSDS File Delivery Protocol (CFDP), for missions that require this type of product to make it easier for mission operations. However, for Lunar IceCube and other EM-1 CubeSats, the mission interface is at the frame level, under the Return All Frames or Return Channel Frames service.

For radiometric data, the carrier phase and ranging phase measurements of both uplink and downlink are relayed to JPL via the DCD. The information is packed into JPL-specific data format and delivered in real time to mission operation system. The same information can also be packaged in the CCSDS Track Data Message (TDM) format, for maximum interoperability with users who use TDM format. From these data products, the mission navigation team can compute the observed Doppler and ranging, which then helps them with orbit determination for spacecraft navigation

For command data, mission operation system will directly connect to the Uplink equipment at Morehead, using the CCSDS Space Link Extension Forward Command Link Transmission Unit (CLTU) interface specification. This SLE interface also enables user authentication and allows the mission operation team to control the radiation of commands to their spacecraft.

To enable data delivery between mission users and their respective spacecraft, many aspects of service management need to be considered. Service management refers to the antenna scheduling, generation of spacecraft prediction data, equipment monitoring and configuration management. Several service management interfaces are reflected in Figure 2.

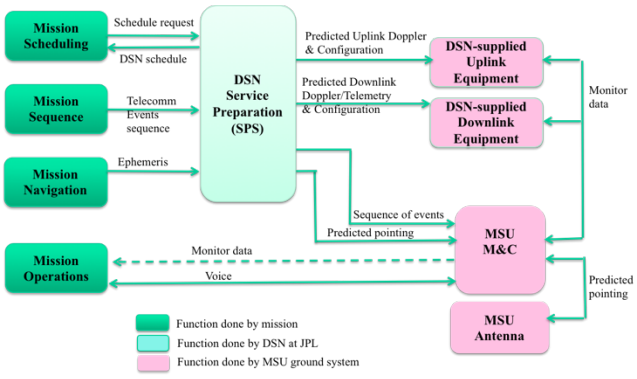


Figure 2. Service management data flow.

(1) Antenna Scheduling – The MSU antenna is schedulable by the DSN Scheduling. This allows for integrated support on missions, especially those that use both Morehead and DSN antennas. Information on any planned antenna maintenance, or science-related activities such as radio astronomy observations, will be inputted by the MSU team into the DSN schedule to indicate the time when the antenna is not available. The remaining time is open for spacecraft tracking, freely assigned by the DSN Scheduling. Through a web browser or application program interfaces, the operation team at Morehead would be aware of upcoming tracks and make necessary preparation for them.

(2) Signal Predictions – Once a track is scheduled, predictions on the expected Doppler frequency, antenna pointing and signal conditions (e.g., data rates, signal power, coding and modulation scheme, times of signal arrival and exit) are generated by the Service Preparation Subsystem (SPS) at JPL. The required inputs for this processing are spacecraft ephemeris and expected spacecraft communications configuration (e.g., planned data rate, coding/modulation scheme, start and stop time of signal acquisition). They are submitted by the mission operation team for their respective spacecraft. These prediction data products are automatically pulled by the Monitor/Control (M&C) processing at Morehead via a REST (Representative State Transfer) query based on the tracking schedule. The data are then distributed to appropriate equipment at Morehead - predicted pointing go to the antenna controller; expected downlink frequencies with imbedded Doppler and the expected signal conditions are given to the receiver; coding configuration is given to the decoder; uplink frequencies are given to the uplink controller, along with other configuration information necessary for command and ranging signal generation. The sequence of events for a given pass, generated by the SPS, would also indicate the start and stop time of the track, as well as any configuration changes (e.g., date rate) in mid track. With this information, the Monitor & Control at the MSU can automatically configure the equipment for the track and reconfigure the link to accommodate any subsequent changes in mid pass.

(3) Monitor data – Monitor data from all equipment is collected by the MSU M&C and presented to the MSU operators – either an operating staff member or students. Monitor data from the DSN-provided equipment, because of the way it was built to work with the DSN monitor control infrastructure, which is not deployed at Morehead due to implementation constraints, requires a special application program interface to be built to route the data to the MSU M&C. At this point, monitor data is locally stored at the MSU, rather than being routed to the DSOC and subsequently provided to MOC. In the case of anomalous event, monitor data can be extracted and delivered to the mission operation team for post-pass diagnostics.

Voice communications link – The MSU team will be able to exchange information with the DSN operations and mission operation teams via the standard DSN operational voice networks. The voice equipment supplied at MSU is

 ISBN: 978-1-61208-694-1

the same as those used in the DSN operations center. Voice-over-internet-protocol data are flown over the same communications link that supports telemetry, command and radiometric data delivery.

IV. SYSTEM PERFORMANCE

Table 1 shows the improvement in the system capability at X-band, before and after this upgrade. By employing a cryogenic low noise amplifier, the system noise temperature is expected to drop by a factor of two, from 215 K to under 100K, yielding a higher signal to noise ratio and making signal detection much easier. Coupled with the use of more complex but highly efficient forward error coding such as Turbo or Low Density Parity Check (LDPC) codes, the system can operate at a much lower signal power threshold, within 1 dB of the Shannon limit of information channel capacity.

TABLE I. PRE AND POST UPGRADE PERFORMANCE

Performance Measure	Pre-Upgrade	Post-Upgrade
X-Band Frequency Range	7.0 – 7.8 GHz	7.0 – 8.5 GHz
LNA Temperature	70 K	< 20 K
System Noise Temperature	215 K	<100 K
Antenna Gain	62 dBi (@7.7 GHz)	62.7 dBi (@8.4 GHz)
System Noise Spectral Density	-175 dBm/Hz	<-178 dBm/Hz
G/T at 5° Elevation	37.5 dB/K	40.4 dB/K
Time Standard	GPS (40 ns)	Hydrogen maser (1 ns/day)
EIRP	N/A	93.7 dBW
HPBW	0.124 deg	0.115 deg
SLE Compliance	N/A	Yes
CCSDS Compliance	N/A	Yes
Forward Error Coding	Reed Solomon/Convolutional	Reed Solomon/Convolutional, Turbo, Low Density Parity Check
Radiometric	Angle, Doppler	Angle, Doppler, Ranging

The use of the Hydrogen MASER significantly improves the timing accuracy, by an order of magnitude compared to the previous reference to the Global Position System. This in turn improves the accuracy of radiometric data (i.e., reducing the noise in Doppler and ranging measurements), as well as telemetry data time tag.

Two significant features being added to the system with this upgrade are the ability to uplink at X-band and to conduct ranging. For ranging measurements, both DSN-specific sequential ranging [3] and pseudo-noise ranging [4] are supported. A full CCSDS-compliant pseudo-noise ranging is expected to be available in 2019.

V. SPACECRAFT TESTING

Because the collecting aperture of the 21-m antenna at Morehead is about 40% that of the 34-m antennas in the DSN, and the system noise temperature is about 4 times higher, there is about 10 dB difference in the received G/T. Thus, only a subset of X-band missions currently supported by the DSN that have sufficiently large link margin, above 10 dB, can be tracked by the Morehead antenna on a non-interfering basis, i.e., without requiring spacecraft to make any change in its downlink data rate. MAVEN, a Mars orbiter currently at low data rate of 23 symbols/s, provides a perfect opportunity for Morehead antenna to track it and be able to decode telemetry data. Osiris-Rex and Hayabusa2 signals, due to the use of higher data rate, leave little link margin to allow decoder lock up. Thus, Morehead tracking of Osiris Rex and Hayabusa2 can achieve carrier demodulation and symbol synchronization, but would not be able to decode the telemetry data.

Figure 3 shows the received carrier SNR (Pc/No) and symbol SNR (SSNR) from Hayabusa2 at both Morehead 21-m (designated as Deep Space Station DSS-17) and DSN 34-m antenna (DSS-25 at Goldstone, California). Both measurements indicate a G/T difference of 11.3 dB.

For Osiris Rex, Figure 4 shows a difference of 10.3 dB between the two sites (Morehead DSS-17 and Goldstone DSS-24), based on PcNo measurements.

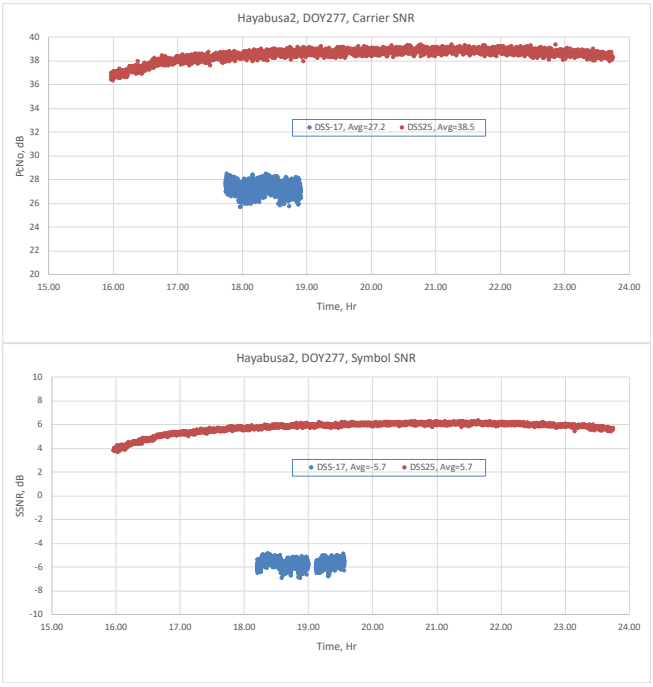


Figure 3. Hayabusa-2 Carrier and symbol SNR

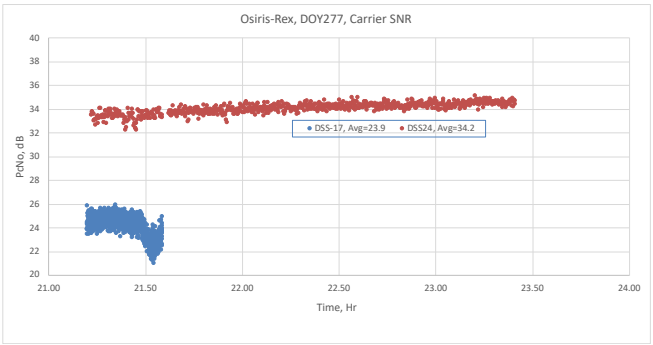


Figure 4. Osiris Carrier SNR

Similarly, Figure 5 shows measurement with MAVEN spacecraft. The PcNo difference was about 11.0 dB; however, due to unknown cause, we were not able to achieve subcarrier lock at Morehead station. As a result, the symbol synchronization suffered excessive degradation,

 ISBN: 978-1-61208-694-1

resulted in a 17 dB difference between Morehead and DSN antennas.

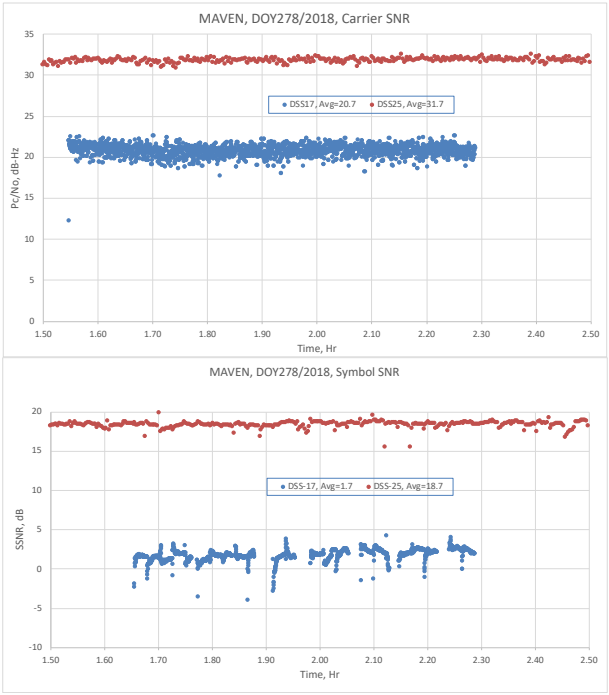


Figure 5. MAVEN Carrier and symbol SNR

Given the difficulty with achieving subcarrier lock and symbol synchronization with Osiris-Rex and MAVEN, we plan to repeat these tests in the near future to collect more data on the performance and to demonstrate a complete data processing all the way to telemetry frame decoding. We also plan to conduct testing with MarCO spacecraft [5], with support from the mission to reduce the data rate down to the level appropriate for the Morehead antenna for a full decoding.

One may notice that in all three tests, the data segment of Morehead was much shorter than the DSN track. This is due to limited test time available, because of ongoing support to the Asteria mission.

VI. CONCLUSION

In summary, this paper describes a tracking ground station being developed at the Morehead State University that optimally combines commercial products and specialized equipment developed for the Deep Space Network. The hybrid architecture will result in a low-cost implementation, with maximum interoperability with DSN antennas and compliance with the CCSDS specifications. Upon completion of the upgrade, the Morehead 21-m antenna system will have full operational telemetry, tracking and command capability at X-band. The system can help to offset some of DSN tracking load in time of heavy demand and is particularly applicable to future CubeSat missions due to the university's strong involvement with CubeSat communities. The system is undergoing testing with spacecraft currently in operation. Preliminary results show the system is functional with successful carrier demodulation and symbol synchronization for spacecraft under test. The G/T measurement indicated Morehead antenna is 10.3 – 11.3 dB lower than that of the DSN 34-m antenna; however more tests will be required to establish more accurate performance and to ensure that we can successful decode telemetry frame from Maven as expected from the link analysis.

ACKNOWLEDGMENT

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Special thanks to J. Yang at JPL for technical assistance with signal acquisition, and to W. Chen and G. Hewitt at Peraton Corporation for assistance with antenna scheduling.

REFERENCES

[1] B. Malphrus, "University SmallSat Programs and Technical Issues", Keck Institute for Space Studies Workshop on New Approaches to Lunar Ice Detection and Mapping, July 2013, http://kiss.caltech.edu/workshops/lunar_ice/presentations1/malphrus2.pdf, retrieved February, 2019.

[2] M. Smith, et al., "On-Orbit Results and Lessons Learned from the ASTERIA Space Telescope Mission", 32th Annual AIAA/USU Conference on Small Satellites, Logan, UT, August 2018, https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=4067&context=smallsat, retrieved February, 2019.

[3] P. Kinman, Sequential Ranging, 810-005, Rev. E, Telecommunications Link Design Handbook, https://deepspace.jpl.nasa.gov/files/810-005/203/203C.pdf, retrieved February, 2019.

[4] P. Kinman, Pseudo-noise and Regenerative Ranging, 810-005, Rev. E, Telecommunications Link Design Handbook, https://deepspace.jpl.nasa.gov/files/810-005/214/214A.pdf, retrieved February, 2019.

[5] Mars CubeSat One, Jet Propulsion Laboratory CubeSat, https://www.jpl.nasa.gov/cubesat/missions/marco.php, retrieved February, 2019.

 ISBN: 978-1-61208-694-1

Satellite 5G: IoT Use Case for Rural Areas Applications

Sastri Kota

University of Oulu,
Oulu, Finland
Email: sastri.kota@gmail.com

Giovanni Giambene

CNIT - Dept. Information Eng. and Math. Sciences, University of Siena,
Siena, Italy,
Email: giambene@unisi.it

***Abstract*—One of the key drivers for next-generation mobile communications, 5G, is the support of the Internet of Things (IoT) with billions of objects being connected to the Internet with low latency. The 5G technology will support the realization of smart cities, smart environments, and big data applications. Within the 5G framework, the terrestrial services can be augmented with the recent development of High Throughput Satellite (HTS) systems and mega-Low Earth Orbit (LEO) satellite constellations. In this paper, we investigate the integration of 5G technology and IoT by means of an aerial component composed of drones and satellites for a rural scenario. The prospected system can provide enhanced services, e.g., fire alarm detection, smart agriculture, animal tracking, and plant disease control. A use case of an agriculture application consisting of a number of areas whose sensor data are collected via drones is described. The proposed architecture consists of the drones connected to a satellite system to provide the necessary network control and connectivity. Subsequently, the satellite segment is connected to the terrestrial network and then to the cloud. In this study, we refer to a rural area scenario where drones are used to detect fire alarms collecting sensors data on the field, aggregating them, and then delivering messages via satellite to a control center. An analytical model has been developed to characterize the distribution of the time to detect and deliver an alarm. This study depends on many parameters; in particular, we have investigated the impact of the area size served by a drone, the maximum sensors range, and the sensor duty cycle.**

Keywords–5G; Satellite Networks; UAVs.

I. INTRODUCTION

Recent studies estimate that about 4 billion people still lack Internet access [1]. The cost of a pure terrestrial coverage will quickly become unbearable with the increasing capacity needs for rural, remote, and urban areas. Moreover, terrestrial networks cannot guarantee the access to the Internet to passengers on aircrafts or high-speed trains, as well as users on vehicles on highways or in the countryside. Under these challenging operational conditions, the terrestrial infrastructure has to be complemented by the satellite segment as envisaged by 5G communication systems. Satellites will also support machine-type communications, paving the way to new applications, ranging from smart agriculture, environmental protection, transportation, animal tracking, etc. The new 5G system will be an umbrella system, enabling different Radio Access Networks (RANs) to operate together, including terrestrial base stations [now called g-Node Bs (gNBs)], aerial platforms of different types, including drones and satellites [2].

It is commonly assumed that 5G systems must address several challenges, including higher capacity, higher data rate, lower end-to-end latency, massive device connectivity, reduced cost and consistent Quality of Experience (QoE) provisioning [3]. ITU-R M.2083 Recommendation classifies three different 5G scenarios, as Enhanced Mobile BroadBand (eMBB), massive Machine-Type Communications (mMTC), and Ultra-Reliable Low-Latency Communications (URLLC) [4]. The satellite systems can support these scenarios as follows:

1) **eMBB**: Users in under-served areas, passengers on board vessels or aircrafts, disaster relief 5G services, emergency communications, media, and entertainment content broadcasts, passengers on board public transport vehicles, etc. These applications can be supported by satellite systems at different altitudes, such as Low Earth Orbit (LEO), Medium Earth Orbit (MEO), and GEostationary Orbit (GEO).
2) **mMTC**: Global continuity of service for telematic applications based on a group of sensors/actuators. This scenario is more suitable for lower orbit satellites, like LEO constellations.
3) **URLLC**: Satellite systems (referring here mainly to LEO cases) can support URLLC-like services that require high reliability and high availability but that do not need extremely-low latency because of the large propagation delays.

This paper deals with, on the one hand, the system characteristics of the aerial component of future 5G systems and, on the other hand, with a use case of the mMTC type for rural areas monitoring, fire alarm, pollution detection, etc. This paper has been developed within the framework of the 5G satellite working group of the 5G IEEE Roadmap initiative [5].

After this introduction, this paper is organized as follows: Section II provides a survey on the state of the art of multi-layer architectures explaining the originality of this work; Section III deals with the aerial component (i.e., drones and satellites) of 5G systems; a system architecture is described in Section IV; sensor technologies are detailed in Section V; our case study dealing with the monitoring of rural areas by means of sensors connected via drones and satellites is provided in Section VI; finally, Section VII draws concluding remarks.

II. STATE OF THE ART

In view of future 5G systems, a multi-layer architecture is envisaged where low altitude drones, Unmanned Aerial Vehicles (UAVs) and High Altitude Platforms (HAPs) can be jointly used to provide a focused coverage or coverage extension to 5G systems with the support of a satellite component with LEO and/or GEO satellites. Layers typically correspond to the

 ISBN: 978-1-61208-694-1

different altitude levels, where a higher altitude implies a wider coverage and a larger 'responsibility' in network management functions and routing support. The scenario of a multi-layer network with drones cooperating with satellites is a relatively new system concept that has received increasing interests recently. Li et al. [6] provide a very general overview on the use of UAVs and their role in an integrated network where there is also the possibility of UAV-to-satellite communications. Shi et al. [7] envisage a multi-layer architecture including a terrestrial part and the aerial component made of both drones and a network of LEO satellites. The interest of this study is on how to interconnect the layers optimally. This study is quite general and does not address the sensor scenario for rural areas and the need of providing services with low latency where there is not a terrestrial infrastructure to interconnect to the Internet.

Huo et al. [8] study a multi-layer aerial component with interesting details for the design of the UAV segment, but there is no consideration of the services, the design of the aerial component, and the impact of the latency experienced by alarm notifications. Finally, standardization bodies [like the European Telecommunications Standards Institute (ETSI), the Third Generation Partnership Project (3GPP), and the International Telecommunication Union - Radiocommunication Sector (ITU-R)] are also interested in the aerial segment with drones, HAPs, and satellites cooperating with the terrestrial 5G segment, as shown in [9] where scenarios and architectures are addressed.

The original contribution of this work is to build on the basis of the above architectures a feasibility study for a scenario where the data provided by sensors are collected by drones and sent via satellite in remote areas with no terrestrial Internet connectivity. A modellization effort has been pursued to characterize the service provided via drones.

III. The Aerial Component of 5G Systems

We provide below an introduction to the different technologies and related networks for the aerial component of 5G systems.

A. Satellites

A High Throughput Satellite (HTS) has many times the throughput of a traditional Fixed Satellite System (FSS) for the same amount of allocated frequency in orbit. These satellites take advantage of high frequency reuse and multiple spot-beams to increase the throughput. A typical HTS satellite can have a capacity of hundreds of Gbit/s. New HTSs typically provide download speeds of more than 10 Mbit/s per user.

The different beams of an HTS satellite reuse the bandwidth according to a typical 4-color reuse pattern (two frequency slots and two polarizations). The satellite capacity can be increased if the different antenna beams of the satellite can use the same frequency band (i.e., full frequency reuse). This approach causes significant interference on a beam because of adjacent beams. A possible solution to reduce the inter-beam interference is adopting a precoding scheme (at the gateway, forward path) where the different transmitted signals on the distinct beams are multiplied by suitable coefficients aimed to orthogonalize them. To do so, an accurate channel estimation is needed so that it is possible to compute the coefficients of the precoding matrix used at the sender to compensate for the interference [10].

HTS platforms have been designed to serve the consumer broadband broadcast market; however, some of them are also offering services to government and enterprise markets, as well as to terrestrial cellular network operators experiencing a growing demand for broadband backhaul to rural cell sites. For instance, **ViaSat-2** [11] is a commercial-communication GEO HTS with a throughput of 300 Gbit/s. This satellite will provide satellite Internet to North America, parts of South America, including Mexico and the Caribbean, and to air and maritime routes across the Atlantic Ocean.

The propagation delay from a GEO satellite to the earth is about 256 ms and that from a MEO or LEO with an altitude lower than 10,000 km is in the range 10 - 70 ms, which is comparatively shorter, but still not negligible. This is why researchers have been more interested in MEO and LEO satellites in recent years. These non-GEO systems have global or quasi-global coverage, but many satellites (i.e., a constellation) are needed to cover all the earth. We consider here mega-LEO satellite constellations that are being developed with services foreseen by 2020. The proposed systems aim to provide access to the Internet with a quality comparable to that of terrestrial systems. The satellite segment comprises many satellites and several terrestrial GateWays (GWs) that are interconnected to the Internet. A dedicated terrestrial network is also used to interconnect the GWs. In some cases, there are inter-satellite links to allow the direct exchange of data among neighbor satellites and to perform routing in the sky. The frequencies currently adopted are in Ku and Ka bands. Satellite systems will also exploit higher frequency bands, such as Q/V/W.

We can consider the following examples of mega-LEO satellite constellations:

LeoSat [12] foresees a constellation (2022) of 78-108 high-throughput Ka-band satellites in LEO polar circular orbits at an altitude of approximately 1,400 km. These satellites form a high-throughput mesh network interconnected through laser inter-satellite links. A ground-based Virtual Private Network (VPN) interconnects the GWs with a public data network. One terminal can use a bandwidth of up to 500 MHz on both uplink and downlink.

Moreover, the **OneWeb** [13] system consists of a constellation of 720 LEO satellites in near-polar circular orbits at an altitude of 1,200 km. OneWeb will provide the users with high speed up to 50 Mbit/s and low latency lower than 50 ms and plans to interoperate with terrestrial mobile operators. It is expected that approximately 50 or more GW earth station sites will be deployed over the time.

The SpaceX satellite system, called **Starlink** [14], consists of two sub-constellations of satellites. A first LEO constellation is composed of 4,425 satellites and operates in Ku and Ka bands at altitudes around 1,110 km to provide a wide range of broadband communication services for residential, commercial, institutional, governmental, and professional users worldwide. The second component of Starlink will be based on another LEO constellation operating in the V band, comprising 7,518 satellites at altitudes around 340 km.

B. UAVs

UAVs are considered here as autonomous communicating nodes. UAVs can be classified into two categories: fixed-wing versus rotary wing. For example, Fixed-Wing UAVs (FW-UAVs) usually have high speed and heavy payload, but they

must maintain a continuous forward motion to remain aloft, thus are not suitable for stationary applications like close inspection. To minimize the Doppler shift and the associated system design challenges, FW-UAVs need to cruise at the slowest possible speed. In contrast, Rotary-Wing UAVs (RW-UAVs), such as quadcopters, though having limited mobility and payload, can move in any direction as well as to stay stationary in the air. Thus, the choice of UAVs critically depends on the applications. UAVs must use dedicated wireless links (mmWaves, free-space optical channels, sub-6 GHz technologies such as LTE) to connect to the core network.

UAVs can also be categorized, based on their altitudes as HAPs and Low Altitude Platform (LAPs) [15] as follows:

1) **HAPs** can be balloons, which are quasi-stationary and operate in the stratosphere at an altitude of approximately 20 km above the earth's surface. HAP-based communications have several advantages over the LAP ones, such as wider coverage and longer endurance. Thus, HAPs are in general preferred for providing reliable wireless coverage for a large geographic area; on the other hand, HAPs are costly.
2) **LAPs** can fly at altitudes of tens of meters up to a few km and can quickly move. LAPs have several important advantages. First, on-demand UAVs are more cost-effective and can be much more swiftly deployed by means of LAPs that are especially suitable for unexpected or limited-duration missions. LAPs can establish short-range Line-of-Sight (LoS) communication links in most scenarios. LAPs can be used for data collection from ground sensors for monitoring purposes.

A balloon-UAV (HAP category) may be the most suitable aircraft for carrying a heavy 5G base station and hovering over the sky for the longest duration. Considering the significant height and coverage it can achieve, an energy-effective balloon-UAV can serve as a 5G terrestrial macrocell base station (gNB). UAV-assisted 5G communications have numerous use cases, including terrestrial base station offloading, swift service recovery after natural disasters, emergency response, rescue and search, and information dissemination.

From an industry perspective, an example of a recent project employing HAPs for wireless connectivity is the Google's Loon project [16]. Moreover, the Facebook's Internet-delivery project via drones has been stopped recently. Finally, Qualcomm and AT&T are planning to deploy UAVs for enabling wide-scale wireless communications in 5G systems.

IV. System Architecture

While Software-Defined Networking (SDN) aims to separate the control plane from the data plane, Network Functions Virtualization (NFV) allows the abstraction of the physical network in terms of a logical network, thus implementing network functions in software. The 5G physical infrastructure consists of the aerial component (the satellite RAN belongs to it), a Terrestrial RAN, and the interconnecting transport network. The logical level (network virtualization) consists of logical nodes such as logical GWs for the Satellite RAN and logical gNBs for the Terrestrial RAN. A controller supports the control plane of physical nodes and an NFV manager coordinates the virtualized functions. Virtualizing some functions of the satellite GWs would improve the flexibility and the reconfigurability in the provision of satellite services. Several virtualization alternatives are possible depending on the distinction between the functions that would remain located in the satellite GW and those that would be moved to the centralized and/or virtualized infrastructure. For transparent satellites, the GW can support the gNB, the Radio Network Controller (RNC), and the virtualized Evolved Packet Core (vEPC) interface. For regenerative satellites, the satellite always involves the gNB while the GW always provides the vEPC interface. The RNC can be either located in the satellite or in the GW [17].

In the NFV context, the adoption of network slicing could facilitate the definition of networks customized for certain traffic types and services. For example, there can be different requirements on functionality (e.g., priority, charging, security, and mobility), differences in performance requirements (e.g., latency, mobility, availability, reliability, and data rates), or distinctions in terms of the users to be served (e.g., public safety users, corporate customers, etc.).

Figure 1 below shows the aerial RAN of an integrated mMTC scenario with interconnections among the different elements as follows: we have interconnections between sensors and UAVs, among UAVs, and between UAVs and the satellite. In this system, we consider that UAVs fly periodically over a certain rural area to be monitored, collect the data provided by sensors spread in the area, and send these data via satellite. An alternative to this approach, not considered in this paper, would be that sensors transmit data to local sinks in fixed positions on the field that act as GWs to the network.

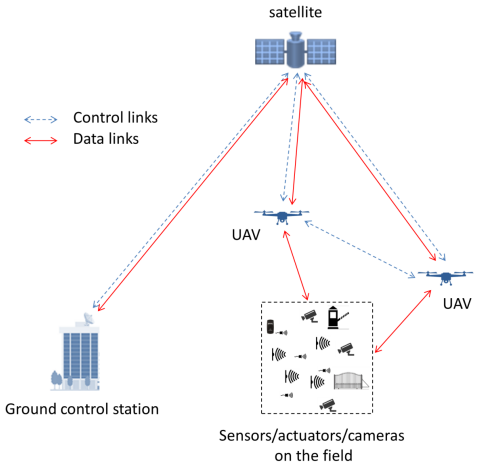


Figure 1. Network architecture for mMTC, aerial RAN.

V. Sensor Technologies

ZigBee [18] is the most popular industry wireless mesh networking standard for connecting sensors, instrumentation and control systems. ZigBee is the classical Internet of Things (IoT) technology. ZigBee is an open, global, packet-based protocol designed to provide an easy-to-use architecture for secure, reliable, low-power wireless networks. ZigBee is a low data rate wireless system based on the IEEE 802.15.4 standard. IEEE 802.15.4 specifies a total of 27 half-duplex channels

 ISBN: 978-1-61208-694-1

across the three frequency bands (868 MHz, 915 MHz, and 2.4 GHz). Channel data rate ranges from 20 kbit/s to 250 kbit/s. The transmission range depending on the frequency band can be from 200 m to 1 km. For instance, the free-space transmission range is around 300 m at 900 MHz (assuming transmission power P_t = +5 dBm, antenna gains G_t = G_r = 1.2 dBi, and received power level of P_r = -105 dBm). Analogously, the maximum transmission range at 2400 MHz is 64 m with similar numerical assumptions as those used for 900 MHz, except for P_t = 0 dBm.

To save the batteries, each node can alternate between awake and sleeping phases. In the awake phase, nodes are active and can communicate messages to neighbors. In the sleeping phase, nodes turn their radios off until the next scheduled wake-up time. The duration of sleeping and active cycles are application-dependent and are set the same for all the nodes. A common duty cycle value is of 10% with wake up time of 10 s.

LoRa (the acronym of Long Range) [19] is a wireless technology specifically designed for long-range, low-power Machine-to-Machine (M2M) and IoT applications. The LoRa range is 5 km for urban and 20 km for rural areas. LoRa is based on the IEEE 802.15.4g standard and operates in the Industrial, Scientific and Medical (ISM) frequency band at 868/915 MHz. A spread spectrum modulation is adopted. LoRa bit-rates depend on the spreading factor; the maximum bit-rate is 50 kbit/s. There is a payload size of 243 bytes for each message. Each LoRa GW can manage up to millions of nodes. The LoRa software is open and its use is free for those who comply with protocol specifications.

Sigfox is a French global network operator that builds wireless networks to connect low-power objects such as electricity meters and smartwatches, which need to send small amounts of data. With Sigfox, a device can transmit up to 140 messages per day. The maximum range is 10 km for urban and 40 km for rural areas. Hence, long distances can be achieved while being very robust against the noise. Messages have a payload size of 12 bytes. Sigfox operates in the ISM band and uses a ultra narrow-band modulation; the maximum bit-rate is from 100 to 600 bit/s, depending on the region.

NB-IoT is a narrow-band technology standardized by the 3rd Generation Partnership Project (3GPP) starting with Release 13. For instance, an LTE operator can deploy NB-IoT inside an LTE carrier. NB-IoT numerology is inherited from LTE. In both downlink and uplink, the channel is divided into 12 subcarriers of 15 kHz. The time domain is divided into time slots, each lasting 0.5 ms and consisting of 7 symbols. Time slots are grouped as follows: two time slots form one subframe (1 ms), 10 subframes form one frame (10 ms). To further improve the coverage, a second numerology with 48 subcarriers of 3.75 kHz is introduced. This numerology is used for the preamble transmission of the random access procedure and optionally for uplink transmissions. In this case, the time slot lasts 2 ms and, for the sake of compatibility, one frame is composed of 5 time slots. The maximum payload size for each message is 1600 bytes. One uplink single-tone (subcarrier) data transmission at 15 kHz provides a physical layer data rate of approximately 20 bit/s when configured with the highest repetition factor (i.e., 128) and the most robust modulation and coding scheme. On the other hand, downlink data transmission achieves a physical layer data rate of 35 bit/s when configured with repetition factor 512 and the most robust modulation and coding scheme. The max data rate is limited to 200 kbit/s for downlink and 20 kbit/s for uplink. NB-IoT can allow a nominal maximum range of 35 km.

Table 1 in the next page provides a comparison among the different sensors technologies.

We can differentiate the sensor-based applications according to the type of data that must be gathered from the field. In particular, we can consider two categories as Event Detection (ED) and Spatial Process Estimation (SPE).

- In the first case, sensors are used to detect an event, for example, a fire in a forest or an earthquake. Every remote device has to measure a quantity, compare with a given threshold and send the binary alarm information. The density of nodes must ensure that the event is promptly detected with a suitable probability of success, while maintaining a low probability of false alarm. In case of a concentrator of alarms (i.e., a local sink collecting data from multiple sensors), the sensors, together with the concentrator, could cooperatively carry out the task of alarm detection.
- In the second case, the sensors aim at estimating a given physical phenomenon (e.g., the atmospheric pressure in a wide area). The main problem is to obtain an estimation of the entire behavior of the spatial process based on the samples taken by sensors placed at random positions. The measurements will be then processed in a distributed way by the nodes or centrally at the supervisor. The estimation error is strictly related to nodes density and the spatial variability of the process.

VI. Case Study: Environmental Monitoring of Large Areas for Agriculture

Every year in Tuscany, a region of Italy, the emergencies for forest fires are repeated regularly, with an almost constant risk during the year and an increase in summer, destroying hundreds of hectares of forest. In Tuscany, more than 800 fire events occur every year with an increasing trend. It is impossible to carry out a direct control by operators in the field given its vastness and the need for continuous monitoring during the day. An automatic radio system is therefore needed to collect a variety of data from the territory to be sent via the Internet to a remote control unit. In the fight against forest fires, the rapid and careful delimitation of the perimeter of the burned areas is fundamental. Therefore, the use of low-cost distributed IoT radio sensors makes it possible to monitor and promptly detect fires, which have to be controlled and resolved in a short time. This technological approach can also be used in other contexts, such as monitoring of landslides, levels of air pollutants in wooded areas, hydrogeological risk, smart agriculture, etc.

The issue is that the sensors must be placed in remote areas where there is no Internet access for several kilometers so that there are problems to convey the data on the field to a remote control center. It is, therefore, necessary to make available a *backhaul interconnection* that can provide adequate capacity even in remote areas. The most suitable sensor technologies today such as LoRa, Sigfox, and NB-IoT do not allow to cover such remote areas. LoRa has a range of up to 20 km with bit rates up to 50 kbit/s. Sigfox has a range of 40 km with bit

 ISBN: 978-1-61208-694-1

TABLE I. COMPARISON OF DIFFERENT WIRELESS SENSOR NETWORK TECHNOLOGIES.

IoT/M2M Technology	Local Range/ Uplink and Downlink Bit-rate	Licensed/ Unlicensed Frequency Band	GW: Availability (Y/N)	GW: Supported Interconnection options (wireless, cable, satellite, etc.) with covered ranges	GW: Traffic capacity (Uplink and Downlink)
LoRa	5 km (urban), 15 km (suburban), 20 km (rural). Data rates up to 50 kbit/s.	Unlicensed spectrum	GW needed	Support Ethernet, 3G, wireless, wired and satellite backhaul	Around 150000 to 1500000 packets per day depending on payload size, symbol rate, coding rate ...
Sigfox	40 km in rural areas and 10 km for urban areas. Data rates up to 600 bit/s	Cellular-like, unlicensed spectrum, ISM	GW needed	Support of Ethernet, 3G, wireless, wired and satellite backhaul	Sigfox GW capacity is three times bigger than LoRa GW capacity
ZigBee (IEEE 802.15.4)	Indoor 10 m; Outdoor 100 m (LoS). Data rate up to 250 kbit/s	Unlicensed spectrum	GW needed	Support of Ethernet, 3G, wireless, wired and satellite backhaul	-
Narrow Band-IoT	Up to 35 km. Data rate up to 200 kbit/s for downlink and 20 kbit/s for uplink	Licensed frequency bands	An LTE cellular system (4G) is used to interconnect with base stations	-	-

rates up to 600 bit/s. NB-IoT has a range of 35 km with an uplink bit-rate up to 20 kbit/s. Therefore, these sensor solutions based on terrestrial infrastructures are not suitable for covering large and remote areas and for supporting high capacity for the transmission (if needed) of images or real-time videos. Therefore, the interest of this paper is to show the possibility to collect sensors data by means of drones with interconnection to the Internet via satellite.

A. Agriculture Monitoring Model

In this section, an application for agricultural monitoring is described that is suitable for large rural areas. The idea is to provide the farmer with periodic and precise monitoring of physical parameters (e.g., temperature, air pressure, humidity, plant illness conditions, etc.) for the real-time control of the plant area as well as for fire alarm. On the basis of the monitoring information, it is possible to increase the quality and amount of production, cut costs, and reduce the pollution caused by weed-killers.

Our system is divided into large areas of $D \times D$ size, each of them being controlled by a drone. In each area, there are many sensors to collect data from the field. A Point Poisson Process is the model typically adopted to characterize the distribution of the sensors on the field for a certain area. This means that the sensors are uniformly distributed in the area considered. The data collected by a drone from multiple sensors can also be aggregated before sending to the satellite.

The sensors can designed with an extremely low duty cycle that is also related to theirs density. During the active mode, each sensor collects measurements and transmits these data to the system via drones that are connected to the Internet via a satellite link [18]. The sensor-to-root data traffic (multipoint-to-point) is predominant. Each drone of the FW-UAV type acts as a mobile GW that has to manage different traffic classes with a scheduler able to share the satellite link capacity taking the Quality of Service (QoS) requirements of the different classes into account.

As shown in Figure 2, we consider that each drone operates over a certain area, collecting sensor data at each pass to be delivered via satellite to a control center. Each drone will use a satellite link not only for flight control but also for sensor data delivery. There can be different types of sensors in the field, such as micro-weather stations, infra-red temperature sensors, sensors to monitor plant diseases, hygrometers, etc. We can consider that the link between drones and satellite has to manage multiple traffic classes, including flight control, remote sensors data, video and/or photo traffic. The sensors traffic can be modeled in a simple way [20][21]. For instance, a sensor for alarms could have an ON-OFF duty cycle, sending an alarm in the ON phase only when a certain measurement threshold is overcome. Otherwise, we could have sensors reporting temperature measurements at regular intervals, from a few seconds to hours depending on the application. Each sensor could send a small measurement packet of max 120 bytes.

B. Delay and Area Size Analysis

We have to design the drone fleet to be able to deliver alarms within a certain predetermined maximum delay. Let d denote the maximum range for the transmission between sensors and drones. Let H denote the drone altitude. Let v_d denote the drone speed assumed to be constant. We consider that a drone can receive the signal from a square area of side W (drone visibility area) that can be characterized as follows:

$$W = 2\sqrt{d^2 - H^2}. \quad (1)$$

For the sake of simplicity, we assume that D/W has an integer value equal to N: $D/W = N$. Of course, $D \geq W$.

 ISBN: 978-1-61208-694-1

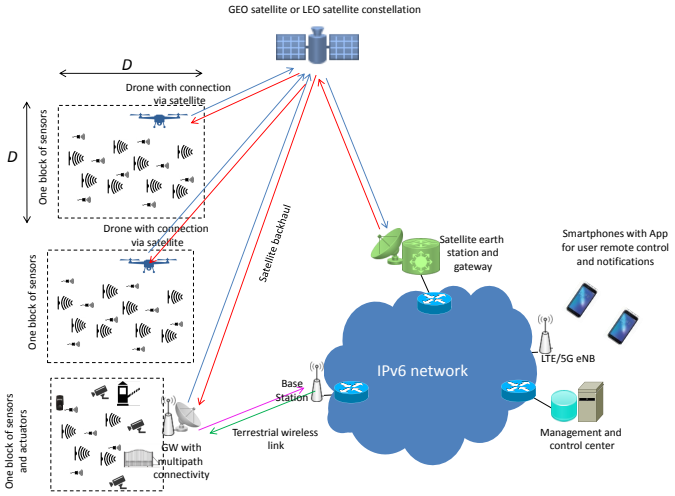


Figure 2. 5G integrated network scenario belonging to the mMTC case that is well suited to represent our sensor-based service.

For the communication between sensors and drones the link budget of the uplink is more critical than that of the downlink. Then, the maximum range d refers to the transmissions from sensors to drones (the transmission power is the one of sensors and the receiver sensitivity refers to the drone). The link budget can be expressed in terms of Allowed Propagation Loss (APL) as:

$$APL = P_t + G_a - S - M, \tag{2}$$

where P_t is the transmission power in dBm of the sensor, G_a denotes the antenna gain in dBi of the receiver (we consider sensors with omnidirectional antennas), S is the sensitivity in dBm of the receiver on the drone, and M represents a margin due to shadowing and interference in dB. Possible values for our scenario are: P_t = 15 dBm, $S = -137$ dBm, $M \approx 10$ dB, G_a = 0 dBi so that the APL becomes equal to 153 dB [22]. The APL term can be converted in terms of distance d (range) using a path loss formula as follows:

$$APL\,(d) = 10\gamma \log_{10}\left(\frac{d}{d_0}\right) + APL\,(d_0) \quad [dB], \tag{3}$$

being γ the path loss exponent and d_0 a reference distance (equal to 1 km). According to [23], the following settings can be adopted for frequencies around 868 MHz (ISM): $\gamma = 2.65$ and $APL\,(d_0)$ = 132.25 dB so that d can be up to 6 km; however, this distance value can be further reduced for lower P_t values and considering noise figures and additional losses in the link budget.

The drone visibility interval t_d can be determined as:

$$t_d v_d = W \rightarrow t_d = \frac{W}{v_d}. \tag{4}$$

The drone can reveal an alarm if it receives the signal during its pass over an active sensor. Each sensor node has an ON/OFF activity to increase the lifetime of its batteries: a sensor can send an alarm only if the alarm conditions are fulfilled during its ON phase. Let T_{ON} (T_{OFF}) denote the mean ON (OFF) phase duration. The drone alarm detection probability due to a sensor P_d can be expressed as follows assuming that $t_d < T_{OFF}$:

$$P_d = \frac{T_{ON} + t_d}{T_{ON} + T_{OFF}}. \tag{5}$$

Let us consider that n independent sensor nodes are present in the alarm area. This parameter n can be related to the density of sensors in the area. Then, the total drone alarm detection probability $P_{d,tot}$ can be expressed as:

$$P_{d,tot} = 1 - (1 - P_d)^n. \tag{6}$$

A drone travels a certain distance P to cover its $D \times D$ service area. The drone cycle time is P/v_d. Since the alarm can occur at any point along this distance, the time that the drone takes to reach the alarm area t_a can be expressed as:

$$t_a = u\frac{P}{v_d}, \tag{7}$$

where u denotes a random variable with uniform distribution between 0 and 1.

Because of the ON-OFF cycle of the sensors, there could be the need for multiple passes over the alarm area to detect the event. The number of passes to detect the alarm is according to a random variable $X \in [1, 2, \ldots.)$ with geometric distribution and parameter $P_{d,tot}$. The first pass needs a time t_a, while the following ones occur after a time P/v_d. Then, the total time to detect the alarm T_{alarm} is a random variable that can be expressed as follows:

$$T_{alarm} = u\frac{P}{v_d} + (X-1)\frac{P}{v_d} + t_{sat} = (u + X - 1)\frac{P}{v_d} + t_{sat}, \tag{8}$$

where t_{sat} denotes the propagation time from the drones via satellite to the terrestrial control station.

The expected alarm notification delay is:

$$E\left[T_{alarm}\right] = \left(\frac{1}{P_{d,tot}} - \frac{1}{2}\right)\frac{P}{v_d} + t_{sat}. \tag{9}$$

We consider that there is a relation between the controlled area size D and the length P of the path of the drone covering that area. The relation between D and P depends on the path selected. We consider the situation depicted in Figure 3.

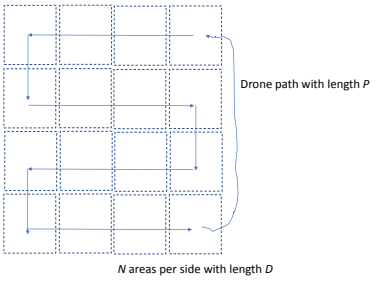


Figure 3. Path of a drone in covering a certain area of size $D \times D$, formed of $N \times N$ squares with side W.

The relation between D and W can be obtained as follows, where we distinguish the sum of horizontal segments and

vertical segments to form the path shown in Figure 3:

$$P = W(N-1)N + 2W(N-1) = \\ = W(N-1)(N+2) = \\ = \frac{D^2}{W} + D - 2W. \quad (10)$$

By substituting all the previous formulas, we obtain the following expression of the mean delay to detect an alarm:

$$E[T_{alarm}] = \left(\frac{1}{P_{d,tot}} - \frac{1}{2}\right)\frac{\frac{D^2}{W} + D - 2W}{v_d} + t_{sat}. \quad (11)$$

The system can be designed to guarantee that the alarm is detected in the 95% of cases within a time limit $T_{max\ delay}$ that is imposed as a design parameter.

$$Prob\{T_{alarm} < T_{max\ delay}\} = 0.95. \quad (12)$$

Since T_{alarm} in (8) depends on two random variables, $u \in (0,1)$ and $X \in \{1,2,3,...\}$, we can use the following approximation:

$$T_{alarm} \approx X\frac{P}{v_d} + t_{sat}, \quad (13)$$

so that the 95-th percentile of T_{alarm} can be expressed by means of the percentile of the geometric distribution of X as:

$$T_{max\ delay} \approx \frac{\ln(1-0.95)}{\ln(1-P_{d,tot})}\frac{P}{v_d} + t_{sat}. \quad (14)$$

Hence, given the $T_{max\ delay}$ value (requirement) the system can be designed by selecting the most suitable *D*, *d*, v_d, n, H, T_{ON}, and T_{OFF} values.

C. Performance Results

Figure 4 shows the behavior of the mean notification delay $E[T_{alarm}]$ as a function of both sensor range d and number N of service areas per D side for n = 2 and 4 alarmed sensors/area. The numerical settings are: v_d = 160 km/h, H = 150 m, T_{ON} = 60 s, T_{OFF} = 600 s, t_{sat} = 250 ms. We can see that the mean notification delay increases with both the sensor range d and the number of service areas N per side.

Figure 5 shows the level curves for constant $T_{max\ delay}$ values depending on the number of sensors n and the duty cycle [$P_{ON} = T_{ON}/(T_{ON} + T_{OFF})$]. We can see that if we need to reduce P_{ON} to increase the battery life, there is the need of a smaller d and then a smaller service area for a drone to keep the same 95-th percentile of the alarm delay.

We can conclude that there is the need of a multi-parameter optimization to select the many system parameters to optimize system costs under requirements in terms of $T_{max\ delay}$.

VII. Conclusions

Many HTS and mega-LEO satellite constellations will deliver Terabits of capacity across the world by 2020-2025. These systems will provide the Satellite RAN of the whole 5G system conceived to be an umbrella system, enabling different technologies to operate together, including UAVs and satellites. In this paper, our major emphasis is on the mMTC scenario for rural areas where sensors on the field are used to monitor fire and pollution events. We have considered a future 5G scenario where drones are used in the territory to collect sensor data

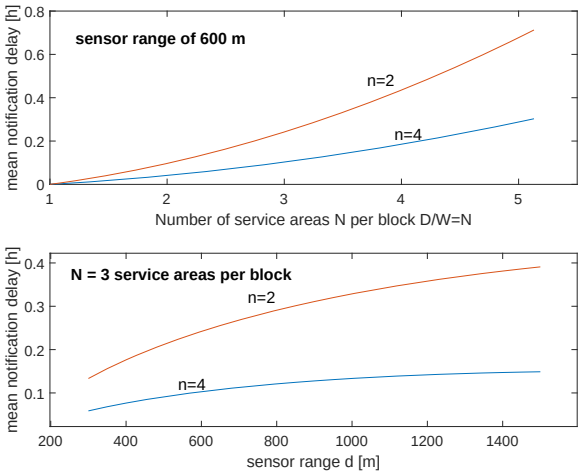


Figure 4. Impact of sensor range d and number N of service areas per block on the mean delay $E[T_{alarm}]$ in hours.

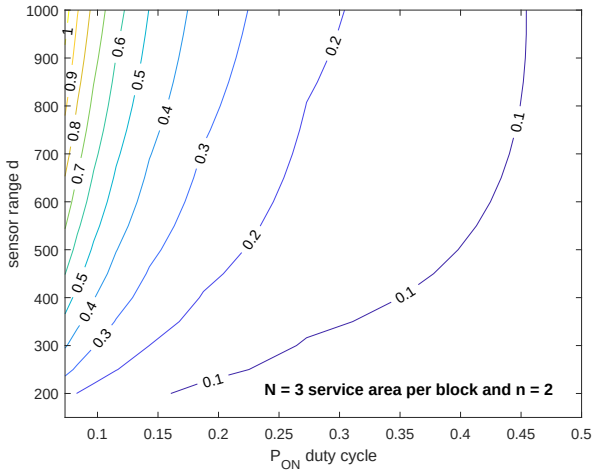


Figure 5. Impact of sensor range d and the duty cycle P_{ON} on the 95-th percentile of the delay, $T_{max\ delay}$ in hours.

that are delivered via a satellite link to a control center. An analytical model has been developed to characterize the time needed to detect an alarm in terms of mean value and 95-th percentile. The model developed in this study can be suitable for a further study to optimize the fleet of drones required to cover the entire rural area under consideration.

References

[1] Report on "World Internet Users Statistics and 2018 World Population Stats", available at the following URL [retrieved: March, 2019]: https://www.internetworldstats.com/stats.htm

[2] G. Giambene, S. Kota, and P. Pillai, "Satellite - 5G Integration: A Network Perspective", *IEEE Network*, pp. 25-31, No. 5, September/October 2018.

[3] A. Gupta and R. K. Jha, "A Survey of 5G Network: Architecture and Emerging Technologies", *IEEE Access*, Vol. 3, pp. 1206-1232, 2015.

[4] 3GPP, "3rd Generation Partnership Project; Technical Specification

Group Radio Access Network; Study on New Radio (NR) to support non-terrestrial networks", TR 38.811V0.2.1 (2017-11) - Release 15.

[5] IEEE 5G and Beyond Technology Roadmap White Paper, available at the following URL [retrieved: February, 2019] : https://futurenetworks.ieee.org/images/files/pdf/ieee-5g-roadmap-white-paper.pdf

[6] B. Li, Z. Fei, and Y. Zhang, "UAV Communications for 5G and Beyond: Recent Advances and Future Trends," *IEEE Internet of Things Journal*, early view available since December 2018.

[7] Y. Shi, J. Liu, Z. Md. Fadlullah, and N. Kato, "Cross-layer Data Delivery in Satellite-Aerial-Terrestrial Communication", *IEEE Wireless Communications Magazine*, Vol. 25, No. 3, pp. 138-143, 2018.

[8] Y. Huo, X. Dong, T. Lu, W. Xu, and M. Yuen, "Distributed and Multi-layer UAV Network for the Next-generation Wireless Communication", Available on arXiv:1805.01534 [eess.SP], 2018.

[9] ETSI, "Integration of Satellite and/or HAPS (High Altitude Platform Station) Systems into 5G and Related Architecture Options", SCN TC-SES DTR/SES-00405 ETSI TR 103 611, December 2018.

[10] A. Guidotti and A. Vanelli-Coralli, "Clustering Strategies for Multicast Precoding in Multi-Beam Satellite Systems", available on arXiv:1804.03891v1 [cs.IT] 11 Apr 2018.

[11] ViaSat-2 system Web site with URL [retrieved: March, 2019]: https://www.viasat.com/viasat-2-service-infographic

[12] LeoSat system Web site with URL [retrieved: March, 2019]: http://leosat.com/technology/

[13] OneWeb system Web site with URL [retrieved: March, 2019]: https://www.oneweb.world/

[14] J. Foust, "SpaceX's Space-Internet Woes", IEEE Spectrum, Vol. 56, No. 1, pp. 50-51, 2019

[15] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems", available on arXiv:1803.00680v1, March 2018.

[16] K. Kamnani and C. Suratkar, "A Review Paper on Google Loon Technique", *International Journal of Research In Science & Engineering*, Vol. 1, No. 1, pp. 167-171, 2015.

[17] 5GCHAMPION EU Project Web site with URL [retrieved: February, 2019]: http://www.5g-champion.eu/

[18] C. Buratti, A. Conti, D. Dardari, and R. Verdone, "An Overview on Wireless Sensor Networks Technology and Evolution", *Sensors (Basel)*, Vol. 9, No. 9, pp. 6869-6896, 2009.

[19] A. Lavric and V. Popa, "Internet of Things and LoRa Low-Power Wide-Area Networks: A Survey", International Symposium on Signals, Circuits and Systems (ISSCS) 2017, July 13-14, 2017, Iasi, Romania.

[20] 3GPP, "Study on RAN Improvement for Machine-Type-Communications", TR 37.868 V11.0.0, September 2012.

[21] A. Sivanathan *et al.*, "Characterizing and Classifying IoT Traffic in Smart Cities and Campuses," *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Atlanta, GA, pp. 559-564, 2017.

[22] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A Survey on 5G Networks for the Internet of Things: Communication Technologies and Challenges," *IEEE Access*, Vol. 6, pp. 3619-3647, 2018.

[23] H. Linka, M. Rademacher, O. G. Aliuy, and K. Jonas, "Path Loss Models for Low-Power Wide-Area Networks: Experimental Results using LoRa", ITG Fachtagung Mobilkommunikation, May 16, 2018.

 ISBN: 978-1-61208-694-1

Receiver Autonomous Integrity Monitoring Performance for Two- Satellites Simultaneous Fault of BeiDou

Ye Ren
National Time Service Center, Chinese Academy of Sciences
Xi'an, Shaanxi, China
e-mail: renye@ntsc.ac.cn

Xiaohui Li
National Time Service Center, Chinese Academy of Sciences
Xi'an, Shaanxi, China
e-mail: xiaohui@ntsc.ac.cn

***Abstract*— As the probability of two-satellites fault simultaneously increase, the difficulty of fault identification in the integrity monitoring become serious and the traditional Receiver Autonomous Integrity Monitoring (RAIM) may not be able to properly deal with this problem. This paper investigates the impact of two-satellites fault at the same time on integrity parameters. Two factors, two-satellites fault mode and range deviation combination, are introduced to simulate the condition of two satellites fault in BeiDou, and the change of integrity parameters (Horizontal Positioning Error (HPE) and Vertical Positioning Error (VPE)) is discussed. The experiment results show that: under two-satellites fault condition, with range deviation increase, HPE and VPE grow. However, HPE/VPE show little change if the faulty satellite mode contains MEO and the range deviation is less than 100 meters.**

Keywords- Two-Satellites fault; HPE/VPE; Kruskal-Wallis test; Multiple Comparison.

I. INTRODUCTION

RAIM is an important integrity monitoring technique for detecting and identifying the satellite faults based on the redundant observations information received by the user receiver. It has the advantages of no other external equipment, low cost, fast detection speed, and realization convenience and so on. It is one of the integrity monitoring algorithms which is widely used at present [1]-[3].

For the GNSS system, as more and more satellites are lunched for positioning and navigation purpose, the possibility of simultaneous two or more satellites fault increases. As for GPS, in the six years between January 1994 and January 2000, there were 0.9 satellites fall in fault [4], that is, the probability of simultaneous of single satellites, two satellites and triple satellites are $1.0274\times10^{-4/}$ hours、1.0556×10^{-8}/ hours、1.0845×10^{-12}/ hours, respectively. Therefore, it is of great practical significance to study RAIM aiming at simultaneous multi-satellites fault. However, the traditional RAIM method based on single satellite fault hypothesis may not be suitable for multi-satellites fault condition. This is because the position error caused by simultaneous multi-satellite fault may fail to follow the normal distribution, so that is the traditional RAIM is applied under simultaneous multi-satellites fault condition, it may reduce the performance of fault detection and fault identification [5][10].

In view of the receiver autonomous integrity monitoring under simultaneous multiple-satellites fault situation, scholars have explored many methods [6]-[10].

The above research works of RAIM for multi-satellites fault are based on the improvement of fault detection and fault identification aspect. For fault detection, the works focus on the error model correction, threshold value determination; for fault identification, they mainly include the construction of test statistics and the search for filtering methods. Most of these studies are based on GPS observations.

With the development of BeiDou constellation, its participation in global satellite navigation (GNSS) activities is increasing, and enough attention should be given to the problem of multiple-satellites fault in the constellation. Due to the unique constellation structure, which includes satellites of GEO (Geostationary Earth Orbit), IGSO (Inclined Geo Synchronous Orbit) and MEO (Medium Earth Orbit), fewer studies have been made on the influence of two or more satellites fault on the receiver autonomous integrity monitoring in BeiDou. Therefore, this paper focuses on BeiDou constellation and discusses in depth the impact of BeiDou two-satellites fault on integrity performance.

The content of the article is arranged as follows: Firstly, two-satellites fault positioning model and integrity parameters are introduced; Secondly, the research plan is proposed and the research method based on non-parametric significance test, which includes Kruskal-Wallis test [12] and multiple comparison, is described. Finally, an experiment is designed to simulate the actual two-satellites fault situation. The final section is a summary of the above analysis.

II. POSITIONING MODEL

A. Positioning model under two-satellites fault

The linearized positioning model with n observations and m estimated parameters:

$$z = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{f} + \boldsymbol{\varepsilon} \tag{1}$$

where z is $n\times1$ measurement vector;

$\boldsymbol{H}$ is n$\times$m is design metric;

$\boldsymbol{X}$ is $m\times1$ state vector;

$\boldsymbol{f} = (0,\cdots, f_i, \cdots, f_j, \cdots, 0)^T$ is fault vector;

$\boldsymbol{\varepsilon} \sim N(0, I_0)$ is $n\times1$ random vector,

I_0 is a diagonal matrix.

If two satellites are faulty at the same time, satellite A

and satellite B correspond to the ith and jth observation separately, then the equation (1) can be divided into two sub-models, which are the fault measurement sub-model and the fault free measurement sub-model, as shown in equation (2):

$$\begin{bmatrix} \boldsymbol{A}_i^T z \\ \boldsymbol{B}_i^T z \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_i^T H \\ \boldsymbol{B}_i^T H \end{bmatrix} x + \begin{bmatrix} \boldsymbol{A}_i^T f \\ \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{A}_i^T v \\ \boldsymbol{B}_i^T v \end{bmatrix} \quad (2)$$

where $\boldsymbol{A}_i = \begin{bmatrix} I_2 \\ 0 \end{bmatrix}$, $\boldsymbol{B}_i = \begin{bmatrix} 0 \\ I_{n-2} \end{bmatrix}$,

I_n is a $n \times n$ Diagonal matrix

Estimation of state vector by least square solution is：

$$\hat{\boldsymbol{X}} = \boldsymbol{S}_0 z \quad (3)$$

where $\boldsymbol{S}_0 = P_0 H^T$, $\boldsymbol{P}_0 = (H^T H)^{-1}$ 。

The state estimation error is：

$$\boldsymbol{v}_z = S_0(\varepsilon + f) \quad (4)$$

$$\boldsymbol{v}_z \sim N(S_0 f, P_0)$$

B. *Position Error*

Position Error (PE) [11], which is defined in equation (5), is one of the most important parameters in integrity monitoring. It represents the accuracy of positioning solution and directly affects the results of subsequent fault detection and fault identification.

$$\Delta\hat{X} = \hat{X} - X \quad (5)$$

where $\hat{X}$ is user estimated position, X is user real position. Under user coordination, $\Delta\hat{X} = (\Delta\hat{X}_E, \Delta\hat{X}_N, \Delta\hat{X}_U)$.

Horizontal Position Error（HPE）along east and north direction is defined as in equation (6)：

$$HPE = \sqrt{\Delta\hat{x}_N^2 + \Delta\hat{x}_E^2} \quad (6)$$

Vertical Position Error（VPE）along up direction is defined as in equation (7):

$$VPE = \sqrt{\Delta\hat{x}_U^2} \quad (7)$$

III. EXPERIMENTS DESCRIPTION

A. *Data Description*

The BeiDou II observed data were collected by a Novatel receiver from May 1st, 2017 to May 10th, 2017， Xi'an. BeiDou II system comprises of 14 satellites with GEO satellite labeled from BD 01to BD 05, IGSO satellite labeled from BD 06 to BD 10, MEO satellite labeled from BD 11 to BD 14. The satellite visibility during the observation period is shown in Figure 1. The statistical result of the visible time for each satellite during the observational period shows that BD01- BD 05 was basically visible throughout the observation day, while the visible time of BD06- BD 10 and BD11-14 are 69% and 23% respectively.

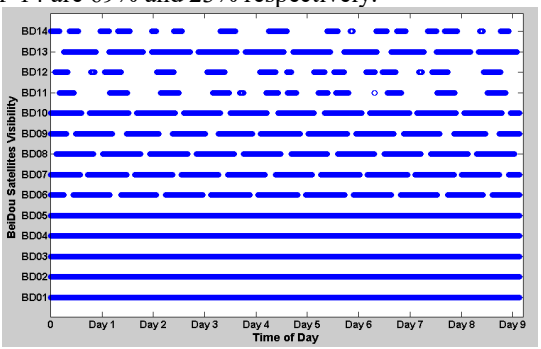


Figure 1. BeiDou Satellite Visibility （1st May, 2017-10th May, 2017，Xi'an）

B. *Experiments Scheme*

Assuming that the system fault is equivalent to some additional range deviations on one or more receiver observations, therefore, two factors are introduced to simulate the situation of two-satellites fault: (1) two-satellites fault mode [A B]. We selected 6 types of modes, as in Table I, and expressed them by [A B], A is the first fault satellite, and B is the second fault satellite; (2) range deviation combination (a, b). As a satellite fault is considered as the sum of range deviation and its corresponding observation measurement, "a" is described as the range deviation on satellite A and "b" is described as the range deviation on satellite B. The range deviation is set increasing from 50 to 200m with 50m of interval. Table II shows the total of 16 sets of (a, b) combinations.

TABLE I. TWO SATELLITES FAULT MODE

[A B]	GEO	IGSO	MEO
GEO	(01，03)	(01，07)	(O1，12)
IGSO	-	(07，09)	(07，12)
MEO	-	-	(12，12)

TABLE II. RANGE DEVIATION COMBINATIONS UNDER TWO SATELLITES FAULT MODE

(a,b)	FAULT SATELLITE B			
FAULT SATELLITE A	(50,50)	(100,50)	(150,50)	(200,50)
	(50,100)	(100,100)	(150,100)	(200,100)
	(50,150)	(100,150)	(150,150)	(200,150)
	(50,200)	(100,200)	(150,200)	(200,200)

 ISBN: 978-1-61208-694-1

IV. SIGNIFICANCE TEST

In order to investigate data feature of position and integrity under two-satellites fault condition, the non-parametric significance test is introduced and it includes two tests, Kruskal-Wallis (K-W) test and multiple comparisons.

A. K-W test

K-W test is one of the classic methods to infer the significant difference among the test sequences and the overall distribution by using sample data in the case of unknown total variance or little knowledge.

There are two hypotheses for the K-W test. Null hypothesis H_0 is the assumption that all sequences come from the same population and the alternative hypothesis H_1 is the assumption that all sequences come from a different population. For k test sequences A_{N_1}, A_{N_2}, ⋯, A_{N_k}, and $N_i = 1,\cdots,k$ is the quantity of ith test sequence. The total quantity of the test sequences is $N = N_1+N_2+\cdots+N_{N_k}$. Then, the data of the k test sequences are merged and sorted in ascending order, and the rank of the variable values is obtained. The k sequences of data are combined and sorted, and the serial number values are given, and the serial number values of the k sequences are respectively summed to obtain R_j (j=1, 2, ..., k), and the test statistics H is established on the basis of equation (8):

$$H = \frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{N_j} - 3(N-1) \tag{8}$$

where $R_{\cdot j} = \sum_{i=1}^{n_j} R_{ij}\,(j = 1,2,\cdots,k)$

After obtaining the test statistics H, the P-value is calculated. The P-value is a probability that the sequences beyond a limitation level under the assumption that the zero hypothesis is correct. It reflects the magnitude of the possibility.

P-value is used to justify whether the hypothesis H_0 is accept or reject by referencing a significant level of 0.05. The significance level refers to the probability that the difference between test sequences caused by sampling error is less than 0.05.

B. Multiple Comparisons

When there is a significant difference among the test sequences based on K-W test, there is reason to consider that all test sequences may not come from the same population. However, this is not to say that there is a difference between any pairs of test sequences. In order to further insight, the difference between the test sequences, multiple comparison is used. For a pair of test sequences A_{N_i} and A_{N_j}, its corresponding confidence interval is constructed [24] in equation (9):

$$\left[\overline{A}_{N_i\cdot} - \overline{A}_{N_j\cdot} - \sqrt{(\frac{1}{N_i}+\frac{1}{N_j})}\sigma \times t_{1-\frac{\alpha}{2}}(f_e), \overline{A}_{N_i\cdot} - \overline{A}_{N_j\cdot} + \sqrt{(\frac{1}{N_i}+\frac{1}{N_j})}\sigma \times t_{1-\frac{\alpha}{2}}(f_e)\right] \tag{9}$$

where $\overline{A}_{N_i\cdot} = T_h / N_h$, $T_h = \sum_{h=1}^{N_i} A_{N_i,h}$, $A_{N_i,h}$ is the hth element in the ith test sequence, α is significant level, $f_e = N - k$, $\sigma^2 = S_e / f_e$ is Error variance, $S_e = \sum_{i=1}^{k}\sum_{j=1}^{N_j}(A_{N_i,j} - \overline{A}_{N_i\cdot})^2$.

The confidence interval obtained by multiple comparison is used to determine whether any two sets of test sequences come from the same population. When the confidence interval contained zero, it indicated that two test sequences may come from the same population. When the estimated interval did not contain zero, it indicated that two test sequences may not come from the same population.

V. EXPERIMENTAL ANALYSIS

The non-parametric significance test is used to discuss the connection between HPE or VPE and the addition of various range deviation combinations under the same two-satellites fault mode.

Under a certain two-satellites fault mode, the HPE and VPE sequences from the original observations are labeled as $HPE_{(a\ \ b)_0}$ and $VPE_{(a\ \ b)_0}$. HPE and VPE sequences from addition of range deviation on the observations are labeled as $\{HPE_{(a,b)_i}\}$ and $\{VPE_{(a,b)_i}\}$, where $(a\ \ b)_i$, $i = 1,\cdots,16$.

The hypothesis of HPE is:

$$H_0: \ \mu_{HPE_{(a\ \ b)_0}} = \mu_{\{HPE_{(a,b)_i}\}}$$

$$H_1: \ \mu_{HPE_{(a\ \ b)_0}} \neq \mu_{\{HPE_{(a,b)_i}\}}$$

The hypothesis of VPE is:

$$H_0: \ \mu_{VPE_{(a\ \ b)_0}} = \mu_{\{VPE_{(a,b)_i}\}}$$

$$H_1: \ \mu_{VPE_{(a\ \ b)_0}} \neq \mu_{\{VPE_{(a,b)_i}\}}$$

In the three-dimensional graphics of Figure 2, the X axis represents 6 types of two-satellites fault modes. Y axis represents the 16 sets of range deviation combinations, and the Z axis represents the P-value obtained by the K-W test with a significant level of 0.05. It shows P-values from all tests are significantly less than 0.05. Therefore, there is a significant difference among HPE sequences with introduction of range deviations. That is to say, under two-satellites fault mode, there are one or more sequences which show significant difference with the addition of range deviations, but we fail to know how many HPE sequences are different. The following multiple comparisons are used to identify the HPE sequences with significant difference.

The analysis on VPE is proceeding in the same way.

 ISBN: 978-1-61208-694-1

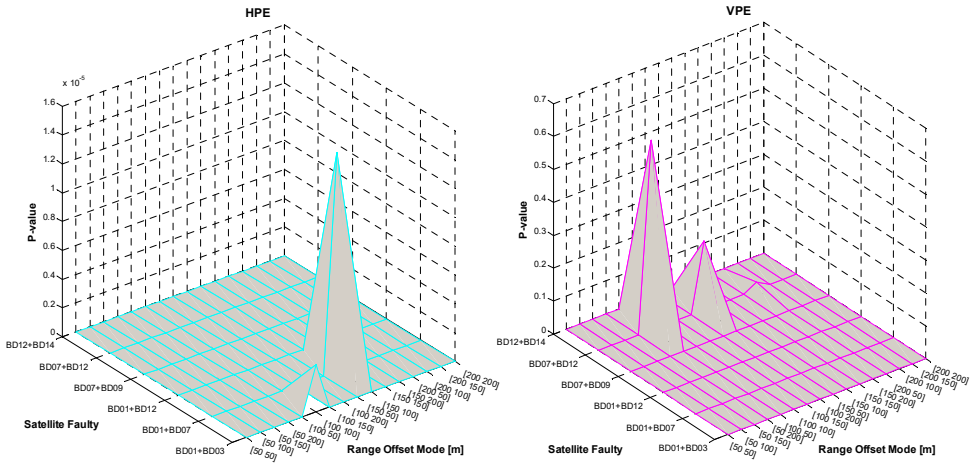


Figure 2. P- values of the K-W test for HPE (left) /VPE (right) under the addition of range deviation combinations and dual-satellite faulty mode (significant level of 0.05)

Table III shows the label scheme in multiple comparisons for HPE. There are 16 comparison groups under each two-satellites fault mode. In each test, $HPE_{(a_0 b_0)}$ from the original observations is compared with each $\{HPE_{(a,b)_i}\}$.

TABLE III. K-W TEST SEQUENCE PAIRING STRATEGY FOR THE INFLUENCE OF DIFFERENT RANGE DEVIATION COMBINATIONS ON HPE（LEFT）/VPE(RIGHT)

	[50 50]	[50 100]	[50 150]	… …	[200 150]	[200 200]
$(a_0 b_0)$	GROUP1	GROUP2	GROUP3	… …	GROUP15	GROUP16

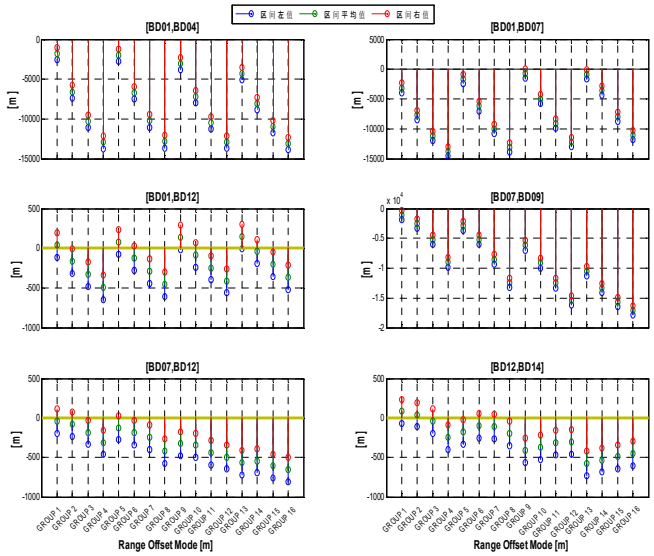


Figure 3. HPE Confidence interval of multiple comparisons before and after introducing range deviation combination.

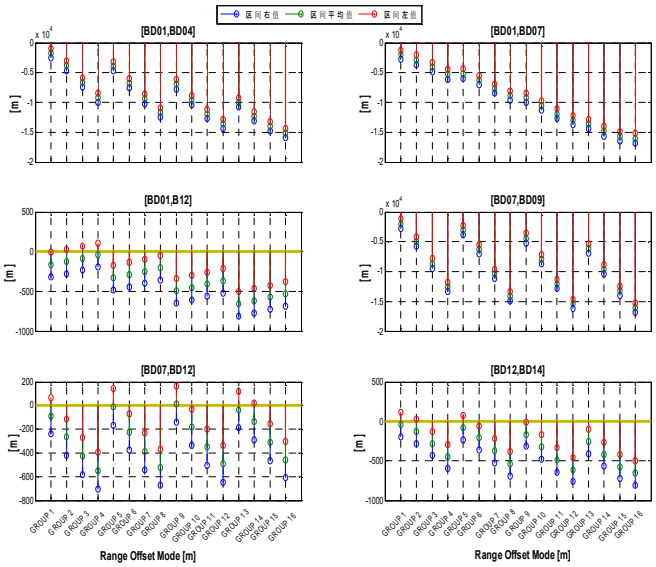


Figure 4. VPE Confidence interval of multiple comparisons before and after introducing range deviation combination.

As shown in Figure 3 and Figure 4：

1) Under three types of two-satellites fault mode, [BD01, BD04], [BD01 BD07] and [BD07 BD09] HPEs show significant difference with addition of range deviation.
2) There is little difference of HPE among [BD01 BD12], [BD07 BD12] and [BD12 BD14]. It is worthy to know that these three types of modes include MEO satellite, such as BD12 OR BD 14.

TABLE IV. RANGE DEVIATION COMBINATIONS FOR TWO-SATELLITES FAULTY MODE WITH NO SIGNIFICANT DIFFERENCE IN MULTIPLE COMPARISONS FOR HPE (LEFT) /VPE (RIGHT)

[BD01 BD12]	[BD07 BD12]	[BD12 BD14]
(50 50)	(50 50)	(50 50)
(100 50)	(50 100)	(50 100)
(100 100)	(100 50)	(50 150)
(150 50)		(100 100)
(150 100)		(100 0)
(200 50)		
(200 0)		

Furthermore, Table IV shows that when the range deviation combination is (50, 50) m, there is hardly any difference among the HPE comparison groups. Specifically, in [BD01 BD12], after the addition of 200m of range deviation on BD01, there is little difference of HPEs against the condition that no additional range deviation. The difference becomes significant when 100m addition of range deviation for BD12.

As for [BD07 BD12], as long as the range deviation is less than 100 meters, there is hardly any difference between the test sequences.

In [BD12 BD14], although both BD12 and BD14 are MEO satellites, it shows that the range deviation combination does not affect them in the same way. That is, when the range deviation combination is (50 100)m, HPE sequences shows a significant difference while HPE sequences have little difference when the range deviation (100 50)m is introduced.

VI. CONCLUSION AND FUTURE WORK

With the development of the Beidou system, the probability of simultaneous failures of two satellites or multiple satellites increases. Based on two-satellites fault condition, this paper discusses the impact of two satellites faults on the integrity parameters HPE/VPE. The result shows that: under two-satellites fault, the range deviation growth leads to the increase of HPE or VPE. However, if the satellite fault mode contains MEO and the range deviation is less than 100m, the effect on HPE/VPE is not significant.

ACKNOWLEDGMENT

This paper is sponsored by the West Talent Foundation of the Chinese Academy of Sciences (XAB2017B09).

REFERENCES

[1] Y. C. Lee, “Analysis of Range and Position Comparison Methods as a Means to Provide GPS Integrity in the User Receiver,” Proceeding of ION 42th Annual Meeting. Seattle, USA, pp.1-4, June 1986.

[2] M. A. Sturza, “Navigation System Integrity Monitoring Redundant Measurement,”Navigation, vol. 35, no. 4, pp. 483-501, 1988

[3] B. W. Parkinson and P. Axeled, “Autonomous GPS Integrity Monitoring Using the Pseudorange Residual,” Navigation, vol. 35, no. 2, pp.255-274, 1988.

[4] S. Feng and W. Ochineg, “User Level Autonomous Integrity Monitoring for Seamless Positioning in All Conditions and Environments,”Proceedings of the European Navigation Conference, Manchester, 2006.

[5] X. H. Xu, C. S. Yang, and R. U. Liu, “Review and prospect of GNSS receiver autonomous integrity monitoring,” Acta

Aeronautica et Astronautica Sinica, vol. 34, no. 3, 2013, pp. 451-463.

[6] Y. C. Lee, "Performance of Receiver Autonomous Integrity Monitoring (RAIM) in the presence of simultaneous multiple satellite faults," Proceedings of Annual Meeting of the Institute of Navigation, pp.687-697, 2004.

[7] S. Hewitson and J. Wang, "GNSS Receiver Autonomous Integrity Monitoring (RAIM) for Multiple Outliers," European Journal of Navigation, vol. 4, no. 34, pp. 47-57, 2006

[8] A. Ene, J. Blanch and J. D. Powell, "Fault detection and elimination for Galileo–GPS vertical guidance," Proc. of the 2007 National Technical Meeting of the Institute of Navigation, pp. 1244–1254, 2007.

[9] J. L. Wang and J. Wang, "Mitigating the effect of multiple outliers on GNSS navigation with M-estimation schemes," International Global Navigation Satellite Systems Society IGNSS Symposiumm, pp.1-11, Dec. 2007.

[10] G. Schroth et al, "Enhancements of the Range Consensus Algorithm (RANCO)," Proceedings of the ION GNSS Conference, Savannah, Georgia, pp.93–103, 2008.

[11] F. Kube, et al. , "Virtual receiver concept and its application to curved aircraft-landing procedures and advanced LEO positioning," GPS Solutions, vol. 22, 41, 2018.

[12] H. W. Lillifors, "On the Kolmogorow-Smirnov test for the exponential distribution with mean unknown," Journal of the American Statistical Association, vol. 64, pp. 387-389, 1969.

[13] M. Hollander and D. A. Wolfe, "Noparametric Statistical Methods". Hoboken, NJ, John Wiley & Sons.Inc, 1999.

 ISBN: 978-1-61208-694-1

Linear Prediction and Rice Codes Based Two-Stage Method for Lossless Telemetry Data Compression

Mohamed A. Elshafey
Computer Engineering Department
Military Technical College
Cairo, Egypt
E-mail: m.shafey@mtc.edu.eg

***Abstract* – The article describes the two-stage lossless compression method for telemetry data, in which the linear prediction method acts as a decorrelator in the first stage, and Rice codes act as an entropy coder in the second stage. A detailed description of implementing linear prediction is presented and applied in experiments on different telemetry frame structures, in the Inter-Range Instrumentation Group (IRIG) standard 106 format. Telemetry data analysis is performed, and compression efficiency is presented based on estimates of the gain in variance and entropy of the output signal from the decorrelator. Based on the experimental results, conclusions are extracted, and recommendations are suggested.**

Keywords—telemetry date frame; Rice coding; Linear Prediction; IRIG-106; lossless compression.

I. INTRODUCTION

Remote sensed data are widely transmitted in the form of telemetry stream to ground stations for analyzing abnormal situations, recovering bad sections in the telemetry stream and other types of post-processing or real-time data processing. Telemetry data has a huge size. So, it should be compressed at the telemetry source before transmission and/or archived in a compressed form [1]-[6].

This paper presents results of studies conducted with the aim of developing an effective two-stage lossless compression method based on linear prediction and Rice codes for different structures of telemetry streams.

The Rice coding algorithm is widely used for lossless telemetry data compression, but its effectiveness is largely determined by the properties of the encoded information. Therefore, it is not suitable in case of highly deviated data [4] [7].

In telemetry systems, the commutator of the data acquisition subsystem measures values from multiple different sensors and outputs a single stream of pulses, which are then modulated and transmitted to a receiving station, where the decommutator returns the serial digital stream back to a parallel form. One complete acquisition cycle of the commutator generates data words, which represent values of each measured parameter. These data words and a synchronization code, which is needed for the decommutation process, establish the telemetry frame.

In a simple commutator, each data word is sampled once per rotation at a rate compatible with the fastest changing measured parameter. In many cases, rate of change of the measured parameters is different, often by several orders of magnitude. Consequently, it is not preferred to perform sampling for the slow changing parameters per each rotation. These parameters can be sampled in a single data word (i.e., a subframe commutation case), e.g., for three slow changing parameters, the main commutator will take three rotations to sample every parameter at least once. Each rotation of the main commutator produces a minor frame, while these three minor frames together form a major telemetry frame. In each minor frame, an additional synchronization code (usually a counter) is added, so the decommutator can distinguish between subframe commutated parameters [1] [9].

In this paper, the process of implementing linear predictions as a first stage before Rice codes is presented with illustrative figures, implemented, tested and verified with real telemetry data in IRIG-106 standard format. This is a comprehensive and open telemetry standard, developed and maintained by the Telemetry Group of the Range Commanders Council (RCC) [8].

The statistical characteristics of the telemetry tested data are examined, and their effects on the prediction efficiency are investigated, and sequentially on the overall compression efficiency. The effect of telemetry frame structure on the selection of the predictor's order and, even, on the suggested number of predictors is determined.

In this paper, a literature survey on recent telemetry compression methods is presented in Section 2, while in Section 3, a survey is provided on successive processes, which are involved in applying linear prediction on telemetry data for lossless compressing. In Section 4, telemetry data preparation and the tested frame structures are described. Experimental tests and analysis for the effects of telemetry data characteristics and frame structure on compression efficiency are demonstrated in Section 5. Finally, in Section 6, conclusions and future works are suggested.

II. LITERATURE SURVEY

The Rice codes method is considered one of the most powerful entropy coders, which is used for telemetry data compression, e.g., it is recommended by The Consultative Committee for Space Data Systems (CCSDS) to compress telemetry data [4].

Many recent researches are involved in the preprocessing layer before entropy coding in order to improve compression efficiency [10]. In [11], a method based on displaying the original data onto geometric surfaces planes is suggested as a preprocessor layer. Another method based on performing XOR operation on subsequent telemetry frames is suggested for preprocessing data before entropy coding [12] [13]. A Neural Network (NN)-based telemetry classification method is suggested in [14].

 ISBN: 978-1-61208-694-1

Prediction based preprocessor layer are widely suggested and used for telemetry compression. Prediction methods based on NN models are illustrated in [15] [16]. An adaptive prediction layer based on the Normalized Least Mean Square (NLMS) algorithm is proposed in [17]. This method depends on the value of the convergence parameter and the smoothing parameter of the NLMS algorithm. It shows results very close and slightly better than those of the non-adaptive prediction method, however, it may be worse for more complex telemetry frame structures (i.e., telemetry frames which contain sub-commutated parameters). As the total length of the major telemetry frame increases, the suggested prediction order of the NLMS based filter should be increased to match the total frame length for a better compression.

III. LINEAR PREDICTION AND RICE CODES BASED TWO-STAGE COMPRESSION METHOD

The effectiveness of any compression algorithm is affected by the statistical characteristics of the data to be compressed and generally, if the compression algorithm can adapt to these characterises, it will generate better results. So, recent lossless compression algorithms are generally described by a series of two stages as shown in Figure 1:

1) Decorrelation stage, which figures out the correlation in the original data to be encoded.

2) Entropy coding stage, in which entropy coding will be performed on decorrelated data, which in general has lower variance and lower entropy value than of the original ones [18]-[24].

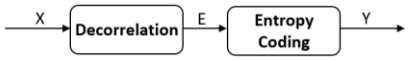


Figure 1. Two stage lossless compression scheme.

A. Rice coding algorithm overview

The efficiency of Rice codes depends on a controllable parameter *k*, which depends on the variance of symbols in the message.

A code word for an encoded number (*v*) by Rice coding algorithm is divided into two parts, i.e., $v=v_i+v_f$, where $v_i = \lfloor v/2^k \rfloor$ and $v_f = v \bmod 2^k$. The coded word is formed by the value of (v_i +1), provided by unary code, and the value of v_f, represented by (*k*) bit binary code.

In this paper, the following formula [23] [25] is used in experiments:

$$k = \log_2(\log_e(2) \times E(|x|)) , \quad (1)$$

where E(|x|) is the average value of symbols of the encoded message.

Rice codes are effective for coding values, which are close to 2^k, and are recommended for the encoding data, whose distribution matches the geometric one.

B. Linear prediction method as a decorrelator

This paper presents a detailed description of applying linear prediction as a decorrelator for reducing correlational dependences in a telemetry stream.

In linear prediction, the value of the sample x_i is predicted by the known values of the previous samples $x_{i-1}, x_{i-2}, \ldots, x_{i-p}$ by the formula:

$$\hat{x}_i = Q\left(\sum_{j=1}^{p} a_j x_{i-j} \right) , \quad (2)$$

where $\hat{x}_i$ is the predicted value of sample x_i, Q is a non-linear operator, which denotes level quantization, a_j is the prediction coefficients and p is the linear prediction order. At each step, the prediction error e_i is calculated by:

$$e_i = x_i - \hat{x}_i . \quad (3)$$

To perform a two-stage lossless compression of telemetry data, the following scheme presented in Figure 2 is implemented, where $\hat{a}_i$ represent the quantized values of prediction coefficients a_i.

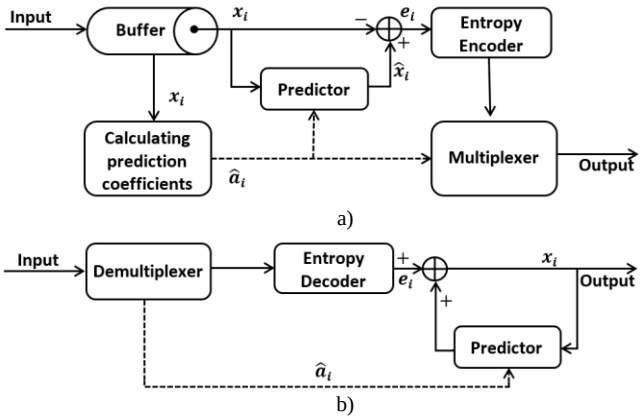


Figure 2. The scheme of two-stage lossless data compression (a) and reconstruction of the original data from the compressed ones (b) [26].

The prediction coefficients are required at the decoder to reconstruct the original signal samples x_i from the prediction error values e_i. So, they are added to the output of the decorrelator.

C. Calculation of predictor coefficient(s)

The optimal values of linear prediction coefficients are obtained by minimizing the Mean Square Error (MSE) of the estimate:

$$e(n) = x(n) - \hat{x}(n) \rightarrow \min (E (e^2(n))) , \quad (4)$$

and based on the principle of orthogonality, thus:

$$E\left[x(n-L) \times e(n) \right] = E\left[x(n-L) \times (x(n) - \hat{x}(n)) \right] = 0 , \quad (5)$$

which can equivalently be written as:

$$\sum_{t=1}^{p} \alpha_t R_{xx}[L-t] = -R_{xx}[L] , \quad (6)$$

where $L=1..p$, $R_{xx}[L]$ is the autocorrelation of x(n), α_t is a vector of length p, $R_{xx}[L-t]$ is a matrix of size $p \times p$ and $R_{xx}[L]$ is a vector of length p. These t equations are Toeplitz

and known as the Yule-Walker equations. A quick and efficient way to solve this matrix is the Levinson-Durbin algorithm, which uses the symmetry of the R_{xx}[L-t] matrix to simplify the calculations from $\{O^3\}$ to $\{O^2\}$ [27].

D. Coding of linear predictor coefficients

Direct quantization of the obtained prediction coefficients has several disadvantages. This operation leads to significant distortions in the frequency domain. Instead, their reflections, using a nonlinear procedure, will be quantized.

Since the values of the prediction coefficients are in the range [–1, + 1], a nonlinear procedure y = arcsin(x), shown in Figure 3, has been implemented [26].

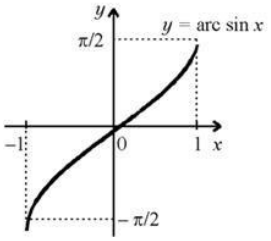


Figure 3. Function diagram of Y=arcsin(X).

The values $\hat{a}_i = \arcsin(a_i)$ are in the range $[-\pi/2, +\pi/2]$. Each value of $\hat{a}_i$ is converted to an 8-bit integer (one bit to represent the sign, and seven bits to represent the output of the arcsin function, which is in the range from 0 to 90). The successive functions, in Figure 4, illustrate the coding process for prediction coefficients for lossless compression.

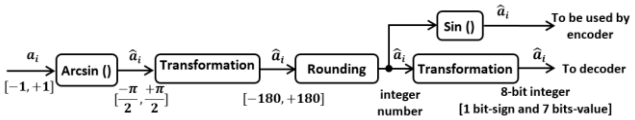


Figure 4. Coding scheme of prediction coefficient using arcsin function.

E. Prediction error transformation

Prediction errors are determined by using (3), which can take both positive and negative values. Therefore, a conversion process is required, which preserves the original range of signal values. The range of error values is equal to the range of values of the input signal to the converter. The error e_i is recalculated into a value δ_i according to the formula:

$$\delta_i = \begin{cases} 2e_i, & 0 \le e_i \le \theta_i, \\ 2|e_i| - 1, & -\theta_i \le e_i < 0, \\ \theta_i + |e_i|, & \text{otherwise}, \end{cases} \quad (7)$$

where $\theta_i = \min(\hat{x}_i - x_{min}, x_{max} - \hat{x}_i)$, x_{min} is the minimum value of the input signal, x_{max} is the maximum value of the input signal. The prediction error is in the range $x_{min} - \hat{x}_i \le e_i \le x_{max} - \hat{x}_i$. The number of bits required to represent errors e_i is the same to represent the signal x_i. If the input signal is represented by positive 8-bit numbers, then $x_{min} = 0$, $x_{max} = 2^8 - 1$.

The inverse transformation is performed according to the following formula:

$$\text{if } \delta_i \le 2\theta_i, \text{ then}$$
$$e_i = \begin{cases} \delta_i/2, & \text{if } \delta_i \text{ even}, \\ -(\delta_i + 1)/2, & \text{if } \delta_i \text{ odd}; \end{cases} \quad (8)$$
$$\text{if } \delta_i > 2\theta_i, \text{ then}$$
$$e_i = \begin{cases} \delta_i - \theta_i, \text{ if } \theta_i = \hat{x}_i - x_{min}, \\ \theta_i - \delta_i, \text{ if } \theta_i = x_{max} - \hat{x}_i. \end{cases}$$

To convert a prediction error to a positive integer value, an error transform block is applied at the output of the decorrelator [28], as shown in Figure 5-a and Figure 5-b at the compressor side and at the decompressor side, respectively.

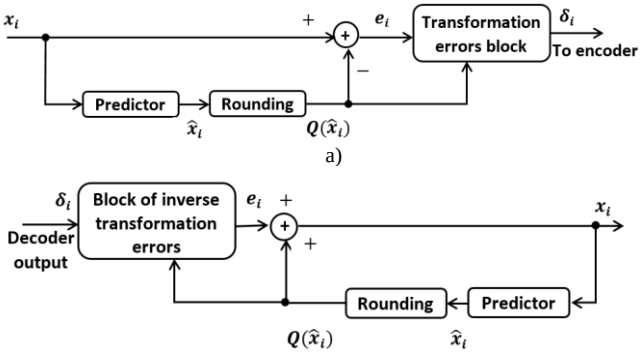


Figure 5. Transformation scheme (a) and inverse transformation scheme (b) of prediction errors.

IV. PREPARING DATA FOR EXPERIMENTS

The telemetry data of automatic control systems have been used in experiments. They are presented in time series of digitized samples of analog sensors, which represent typical telemetry system parameters including temperature, pressure, position data, and so on. These telemetry parameters were obtained in laboratory conditions from real sources. A timed portion of the data samples acquired from nine physical real sensors is shown in Figure 6.

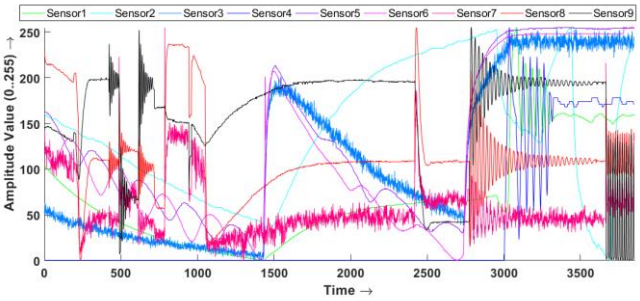


Figure 6. A timed portion of the acquired data samples.

The telemetry data samples used in our experiments are transmitted in frames with a fixed length and a

predetermined fixed internal structure in IRIG-106 standard. In frames, different readings of several sources, i.e., digitized analog sensor readings and readings of digital devices, can be transmitted. Each data source is transmitted in a separate channel of telemetry data recording system in an 8-bit data word in the telemetry frame. The telemetry stream is generated by the telemetry simulation software presented in [29]. The simulator is used to generate different telemetry data streams, based on different telemetry frame structures, in IRIG-106 standard.

Experiments are carried out on two different telemetry frame structures: tf1 and tf2, as shown in Figure 7 and Figure 8, respectively. Telemetry frame structure tf1 consists of 9 channels (DataWords) in one major frame without any subframes, i.e., one level of commutation. An additional piece of information (SYNC_F) of 16 bit (according to IRIG-106 standard) to provide frame synchronization in the telemetry stream, is added.

SYNC_F	DataWord 1	DataWord 2	DataWord 3	DataWord 4	DataWord 5	DataWord 6	DataWord 7	DataWord 8	DataWord 9
SYNC16	Sensor1	Sensor2	Sensor3	Sensor4	Sensor5	Sensor6	Sensor7	Sensor8	Sensor9

Figure 7. Telemetry frame tf1 without subframes.

The second structure tf2 in Figure 9 is a telemetry frame structure, which contains a subframe of length 4, resulting from a second level of commutation attached to the 4th DataWord. A counter as an additional subframe synchronization (SYNC_SF1) is added to each minor frame.

SYNC_F	SYNC_SF1	DataWord 1	DataWord 2	DataWord 3	DataWord 4	DataWord 5	DataWord 6
SYNC16	C1	Sensor1	Sensor2	Sensor3	Sensor5	Sensor4	Sensor7
SYNC16	C2	Sensor1	Sensor2	Sensor3	Sensor6	Sensor4	Sensor7
SYNC16	C3	Sensor1	Sensor2	Sensor3	Sensor8	Sensor4	Sensor7
SYNC16	C4	Sensor1	Sensor2	Sensor3	Sensor9	Sensor4	Sensor7

Figure 8. Telemetry frame tf2 with a subframe at the 4th DataWord.

V. EXPERIMENTAL TESTS AND ANALYSIS

The efficiency of the decorrelation stage can be evaluated by two parameters:

- Prediction gain in variance Prediction_Gain, which is the relation between the variance σ_X^2 of the original data (X) and the variance σ_E^2 of the prediction error (E):

$$\text{Prediction\_Gain} = \sigma_X^2 / \sigma_E^2 . \quad (9)$$

- The entropy H (in bits / symbol):

$$H = -\sum_{i=1}^{Z} P_i \times \log_2(P_i) , \quad (10)$$

where Z is the number of symbols in encoded data and P_i is the probability of each symbol i.

For evaluating the efficiency of the compression algorithm, the compression ratio (R) is checked:

$$R = S_{input} / S_{output} , \quad (11)$$

where S_{input} is the size of the original data (X), and S_{output} is the size of compressed data at the system output (output of entropy coder).

Probability distribution of telemetry data samples is shown in Figure 9. According to this histogram, the values are distributed throughout the scale. The histogram has several pronounced peaks typical of telemetry information, in which the monitored parameters for a long time retain values close to constant, and deviate slightly from them under the influence of noise. It has entropy value $H_x = 7.45$ (bits/symbol) and variance $\sigma_X^2 = 6085.77$.

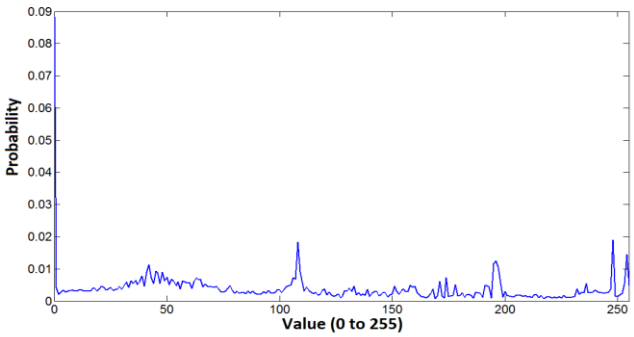


Figure 9. Probability distribution of data samples in telemetry stream.

Applying Rice codes only to the original data (X) leads to a compression ratio R = 1.03 at parameter k = 7, which means nearly no compression.

The linear prediction method is applied on the original data (X), in the first telemetry frame structure tf1. Experimental dependencies of linear prediction efficiency, using (9) and (10), on different prediction orders are shown in Table I.

TABLE I. LINEAR PREDICTION EFFICIENCY AT DIFFERENT PREDICTION ORDERS APPLIED ON TELEMETRY FRAME TF1

Prediction Efficiency	Predictor				
	Order 1	*Order 2*	*Order 8*	*Order 9*	*Order 18*
σ_E^2	5599.01	5565.08	4394.12	78.32	82.42
Prediction_Gain	1.08	1.09	1.38	77.70	73.83
H_E	7.86	7.78	7.71	3.06	3.79

The best results in Table I, i.e., the highest value of prediction gain and the lowest of entropy, are obtained using the 9th order of linear prediction. The probability distribution of the prediction errors of the 9th order predictor is shown in Figure 10, in which the graph has an expressed peak at zero.

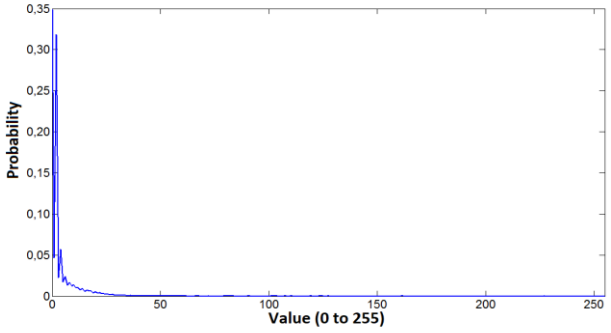


Figure 10. Probability distribution of prediction errors of 9th order predictor applied on telemetry frame tf1.

Applying Rice codes as a second stage, after a 9th order linear predictor at a first stage, leads to a compression ratio value R = 2.13 at parameter k = 3.

The dependences of the compression ratio, (resulting from linear prediction based decorrelation and Rice codes entropy coding), on the order of the linear predictor are shown in Figure 11. We notice that the compression ratio increases dramatically when the order of the linear predictor becomes equal to the count of data words in the repeating structure of the information flow. In this case, this count is equal to the frame length of the main commutator (i.e., order = 9).

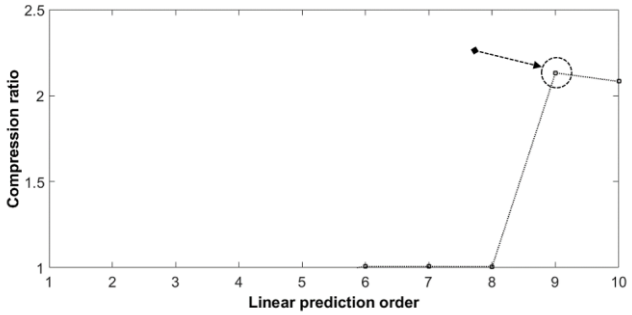


Figure 11. Experimentally obtained dependences of the compression ratio on the order of the predictior for telemetry frame tf1.

Experiments are carried out on the second telemetry frame structure tf2. Based on results obtained, shown in Figure 12, the compression ratio increases explicitly when the predictor order becomes equal to the count of words contained in the four frames of the main commutator (i.e., Best compression ratio R = 1.65 with a linear prediction of order 24).

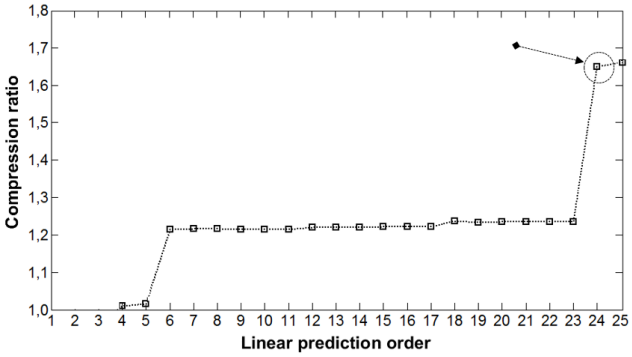


Figure 12. Experimentally obtained dependences of the compression ratio on the order of the predictior for telemetry frame tf2.

From the obtained results, it can be seen that, between readings of the same parameter, located in adjacent frames, there is a strong correlational dependency, and there are usually no such dependencies between adjacent words in the telemetry stream. On the other hand, increasing the order of the predictor to identify these dependencies leads to more complex calculations. In general, the use of the linear prediction method for data decorrelation in a telemetry stream is justified in the case of a small frame length and a simple one level of commutation. If these conditions are not met, then this method of decorrelation is not very efficient or requires significant computational resources.

From the above and based on the fact that the output frames from the main commutator have a fixed length and a constant predetermined internal structure, a linear prediction based decorrelation method is presented. In this method, a decorrelation stage is applied for each telemetry parameter channel, as described in Figure 13. For a given telemetry frame structure in Figure 13-a, a linear prediction is applied on each parameter channel, as shown in Figure 13-b.

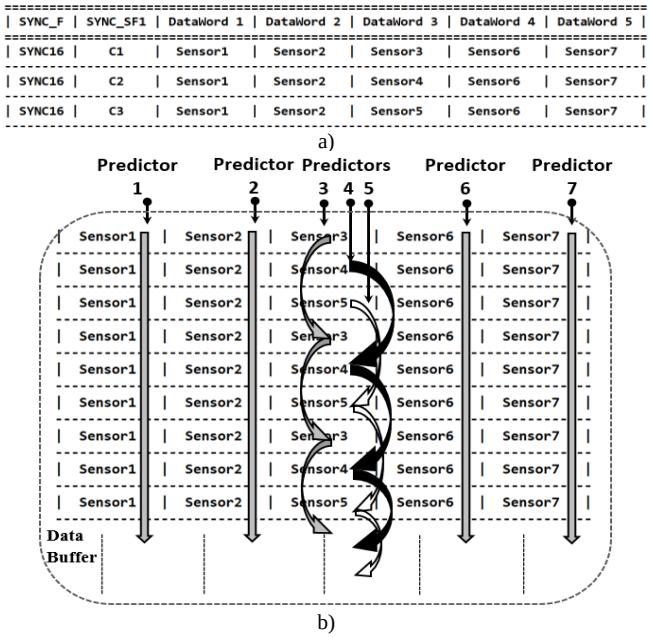


Figure 13. Description of linear prediction of each parameter's channel.

Additional experiments, based on the proposed method, are applied on telemetry frame tf2. It should be noted that the use of first order predictors to predict values of each channel is quite effective.

In experiments, applying 9 linear predictors, each of order 1, for each parameter channel in telemetry frame tf2 leads to a better prediction efficiency than the other obtained from applying a single predictor, as shown in Table II.

TABLE II. LINEAR PREDICTION EFFICIENCY OF DIFFERENT PREDICTORS APPLIED ON TELEMETRY FRAME TF2

Prediction Efficiency	Single linear predictor of order 24 for the stream	First order linear predictor for each parameter channel
Prediction_Gain	28.31	40.84
H_E	4.57	3.34

Applying single first order predictor for each channel, in a two stage Rice based compression method for telemetry frame tf2 achieves compression ratio R=1.96.

 ISBN: 978-1-61208-694-1

VI. CONCLUSION AND FUTURE WORK

Applying linear prediction method as a decorrelator is sufficient enough to transform the probability distribution model of the encoded data samples into a model which is more efficient when implementing entropy coding.

For telemetry major frames without subframes, a single linear predictor of order equal to the length of the major frame (i.e., total number of data words) is an efficient decorrelator, while for telemetry major frames with subframes, applying first order linear predictor for each telemetry parameter channel in the major frame is more effective.

Applying linear prediction for each channel transfers the problem of decorrelation into a parallel one, which can be implemented on Graphics Processing Units (GPUs) for high performance parallel computations.

A preprocessor stage of noise removal from measured telemetry signals can lead to a better efficiency in the subsequent stage (i.e., prediction based decorrelation stage).

Recent deep learning algorithms based on Convolutional Neural Networks (CNN) are promising to provide better and worthy prediction efficiency. As for applying CNNs on onboard systems, a recent method of deep compression should be visited [30].

REFERENCES

[1] F. Carden, R. Jedlicka, and R. Henry, Telemetry Systems Engineering, MA, Norwood: Artech House, 2002.

[2] J. P. Arcangeli, M. Crochemore, J. N. Hourcastagnou, and J. E. Pin, “Compression for an Effective Management of Telemetry Data,” NASA Technical Documents: nasa_techdoc_19940019483, pp. 823-830. 1993.

[3] J. A. Martínez-Heras, D. Evans, and R. Timm, “Housekeeping Telemetry Compression: When, how and why bother,” First International Conference on Advances in Satellite and Space Communications, 2009, pp. 35-40, doi:10.1109/SPACOMM.2009.30.

[4] Lossless Data Compression. Recommendation for Space Data System Standard, Recommended Standard CCSDS 121.0-b-2, Blue Book, May 2012. Available online: https:// public.ccsds.org / publications / BlueBooks.aspx [Retrieved: Feb., 2019].

[5] P. Staudinger et al., “Lossless Compression for Archiving Satellite Telemetry Data,” IEEE Aerospace Conference Proceedings, 2000, vol. 2, pp. 299-304.

[6] D. A. Maluf, P. B. Tran, and D. Tran, “Effective Data Representation and Compression in Ground Data Systems,” IEEE Aerospace Conference, 2008. pp. 1-7.

[7] R. F. Rice, Some Practical Universal Noiseless Coding Techniques - Parts I-III, JPL Tech. Reps. JPL-79-22 (1979), JPL-83-17 (1983), and JPL-91-3 (1991).

[8] Telemetry Standards, IRIG Standard 106-13. Secretariat Range Commanders Council US Army White Sand Missile Range. New Mexico (USA), 2013. Available on-line: http: // www.irig106.org / docs / 106-13 / [Retrieved: Feb., 2019].

[9] S. Horan, Introduction to PCM Telemetering Systems. 2nd Edition. USA: CRC Press, 2002.

[10] A. V. Levenets, “The Basic Principles and Methods of the System Approach to Compression of Telemetry Data,” Journal of Physics: Conference Series, vol. 944, no. 1, IOP Publishing, 2018.

[11] A. V. Levenets, I. V. Bogachev, and E. U. Chye, “Telemetry Data Compression Algorithms Based on Operation of Displaying onto Geometric Surfaces,” In Control and Communications (SIBCON), 2017 International Siberian Conference on, pp. 1-6, IEEE, 2017.

[12] M. A. Elshafey, “A Two-Stage Rice-Based Lossless Compression Method for Telemetry Data,” Journal of Engineering Science and Military Technologies, vol. 2, no. 3, pp. 147-152, doi:10.21608/ejmtc.2018.20314.

[13] D. J. Evans and A. Donati, “The ESA POCKET+ Housekeeping Telemetry Compression Algorithm: Why Make Spacecraft Operations Harder than it already is?” In 2018 SpaceOps Conference, 2018.

[14] X. Shi, Y. Shen, Y. Wang, and L. Bai, “Differential-Clustering Compression Algorithm for Real-Time Aerospace Telemetry Data,” IEEE Access, vol. 6, pp. 57425-57433, 2018, doi:10.1109/ACCESS.2018.2872778

[15] R. Logeswaran and C. Eswaran, “Neural network based lossless coding schemes for telemetry data,” In Geoscience and Remote Sensing Symposium, 1999, IGARSS'99 Proceedings, IEEE 1999 International, vol. 4, pp. 2057-2059.

[16] R. Logeswaran, ”Fast Two-Stage Lempel-Ziv Lossless Numeric Telemetry Data Compression Using a Neural Network Predictor,” Journal of Universal Computer Science, vol. 10, no. 9, pp. 1199-1211, 2004.

[17] M. A. Elshafey and I. M. Sidyakin, “Lossless Compression of Telemetry Information using Adaptive Linear Prediction," Science and Education Bauman MSTU, vol. 14, no. 4, pp. 354-363, Apr. 2014, doi:10.7463/0414.0707364.

[18] N. Fout and K. L. Ma, “An Adaptive Prediction-Based Approach to Lossless Compression of Floating-Point Volume Data,” IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 12, pp. 2295-2304, 2012.

[19] Y. Ren, X. Liu, W. Xu, and W. Zhang, “Multi-Channel Data Structure and Real-Time Compression Algorithm Research”, IEEE Aerospace and Electronic Systems Magazine, vol. 24, no. 11, 2009, pp. 28-35, doi:10.1109/MAES.2009.5344179.

[20] K. Sayood, Lossless Compression Handbook, New York: Academic Press, 2003.

[21] P. D. Johnson, G. A. Harris, and D. C. Hankerson, Introduction to Information Theory and Data Compression, Second Edition. Chapman and Hall CRC, 2003.

[22] G. Mandyam, N. Magotra, and S. D. Stearns, “Lossless Waveform Compression,” In Industrial Electronics Handbook, IEEE/CRC press, 1997.

[23] M. Hans and R.W. Schafer, “Lossless Compression of Digital Audio”, IEEE Signal Processing Magazine, 2001, vol. 18, no. 4, pp. 21-32.

[24] T. Robinson, SHORTEN: Simple Lossless And Near-Lossless Waveform Compression, Technical Report. Cambridge University, Engineering Department, CUED/F-INFENG/TR.156, 1994.

[25] J. H. Kim, Lossless Wideband Audio Compression: Prediction and Transform, PhD dissertation, Institute for Communication Sciences, Technical University Berlin, Germany, 2004.

[26] T. Liebchen, “MPEG-4 Lossless Coding for High-Definition Audio," in Audio Engineering Society Convention 115. Audio Engineering Society, 2003.

[27] M. Ohsmann, “Fast Transforms of Toeplitz Matrices,” Linear Algebra and its Applications, vol. 231, pp. 181-192, 1995.

[28] Lossless Data Compression. Report Concerning Space Data System Standard. Informational Report CCSDS 120.0-G-3 Green Book. April 2013. Available online: https: // public.ccsds.org / publications / GreenBooks.aspx [Retrieved: Feb., 2019].

[29] M. A. Elshafey and I. M. Sidyakin, “Simulation of Telemetry Information Transmission over a Noisy Channel,” Engineering Herald, Electronic Scientific and Technical Journal, Russia, 2014, no. 1, pp. 38-51.

[30] S. Han, H. Mao, and W. J. Dally, “Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding,” International Conference on Learning Representations (ICLR), 2016, arXiv preprint arXiv:1510.00149.

 ISBN: 978-1-61208-694-1

Application of a New Optimal Factorization of the SDRE Method in the Satellite Attitude and Orbit Control System Design with Nonlinear Dynamics

Alessandro Gerling Romero
Department of Orbital Mechanics and Control - DMC
National Institute for Space Research - INPE
Sao Jose dos Campos, Brazil
e-mail: romgerale@yahoo.com.br

Luiz Carlos Gadelha de Souza
Department of Aerospace
Federal University of ABC - FUABC
Sao Bernardo do Campo, Brazil
e-mail: luiz.gadelha@ufanc.edu.br

***Abstract*—The satellite Attitude and Orbit Control System (AOCS) can be designed with success by linear control theory if the satellite has slow angular motions and small attitude maneuver. However, for large and fast maneuvers, the linearized models are not able to represent all the perturbations due to the effects of the nonlinear terms present in the dynamics and in the actuators. Therefore, in such cases, it is expected that nonlinear control techniques yield better performance than the linear control techniques. One candidate technique for the design of AOCS control law under a large maneuver is the State-Dependent Riccati Equation (SDRE). SDRE entails factorization (that is, parameterization) of the nonlinear dynamics into the state vector and the product of a matrix-valued function that depends on the state itself. In doing so, SDRE brings the nonlinear system to a (not unique) linear structure having State-Dependent Coefficient (SDC) matrices and then it minimizes a nonlinear performance index having a quadratic-like structure. The non uniqueness of the SDC matrices creates extra degrees of freedom, which can be used to enhance controller performance; however, it poses challenges since not all SDC matrices fulfill the SDRE requirements. Moreover, regarding the satellite's kinematics, there is a plethora of options, e.g., Euler angles, Gibbs vector, Modified Rodrigues Parameters (MRPs), quaternions, etc. Once again, some kinematics formulations of the AOCS do not fulfill the SDRE requirements. In this paper, we evaluate the factorization options of SDC matrices for the AOCS exploring the requirements of the SDRE technique. Considering a Brazilian National Institute for Space Research (INPE) typical mission, in which the AOCS must stabilize a satellite in three-axis, the application of the SDRE technique equipped with the optimal SDC matrices can yield gains in the missions. The initial results show that MRPs for kinematics provides an optimal SDC matrix.**

Keywords-nonlinear control; SDRE method; satellite control.

I. INTRODUCTION

The design of an AOCS that involves plant uncertainties, large angle maneuvers and fast attitude control following a stringent pointing, requires nonlinear control methods to satisfy performance and robustness requirements.

An example is a typical mission of the INPE in which the AOCS must stabilize a satellite in three-axes so that the optical payload can point to the desired target with few arcsecs of pointing accuracy.

One candidate method for a nonlinear AOCS control law is the State-Dependent Riccati Equation SDRE method, originally proposed by Pearson [11] and then explored in detail by [1]-[3]. SDRE provides an effective algorithm for synthesizing nonlinear feedback control by allowing nonlinearities in the system states while offering great design flexibility through state-dependent weighting matrices. SDRE entails factorization (that is, parameterization) of the nonlinear dynamics into the state vector and the product of a matrix-valued function that depends on the state itself. In doing so, SDRE brings the nonlinear system to a (not unique) linear structure having the SDC matrices and then it minimizes a nonlinear performance index having a quadratic-like structure.

Accordingly, a suboptimal control law is carried out by a real-time solution of an Algebraic Riccati Equation (ARE) using the SDC matrices by means of a numerical algorithm. Therefore, SDRE linearizes the plant about the instantaneous point of operation and produces a constant state-space model of the system. The process is repeated in the next sampling steps, producing and controlling several state dependent linear models out of a nonlinear one.

In this paper, we evaluate the factorization options the SDC matrices for the AOCS exploring the requirements of the SDRE technique. In particular, the non uniqueness of the SDC matrices creates extra degrees of freedom, which can be used to enhance controller performance; however, it poses challenges since not all SDC matrices fulfill the SDRE requirements. Moreover, regarding the satellite's kinematics, there is a plethora of options, e.g., Euler angles, Gibbs vector, MRPs (modified Rodrigues parameters), quaternions, etc. Once again, some kinematics formulations of the AOCS do not fulfill the SDRE requirements.

A good survey of the SDRE method can be found in [1] and its systematic application to deal with a nonlinear plant in [2]. The SDRE method was applied by [4][8]-[10][14] for controlling a nonlinear system similar to the six-degree of freedom satellite model considered in this paper. In [4], a simulator using Euler angles based on commercial software

 ISBN: 978-1-61208-694-1

is defined, whereas [8][9] applied quaternions on commercial software. The application of SDRE method, and, consequently, the ARE problem that arises, have already been studied in the available literature, e.g., [10] investigated the approaches for the ARE solving as well as the resource requirements for such online solving. Finally, [7][12] used SDRE method as a building block for high-level control definition, the former one using sliding model technique and the last using pursuit-evasion game. Recently, [8] proposed the usage of differential algebra to reduce the resource requirements for the real-time implementation of SDRE controllers. In fact, the intensive resource requirements for the online ARE solving is the major drawback of SDRE. Nonetheless, the SDRE method has three major advantages: (a) simplicity, (b) numerical tractability and (c) flexibility for the designer, being comparable to the flexibility in the Linear Quadratic Regulator (LQR) [9]. To the best of our knowledge, since SDC matrices are not unique, there is no work focused on the optimal arrangement of the SDC for the satellite attitude control stabilization. Such optimal arrangement has the potential to increase performance and enhance robustness. Therefore, the first contribution of the present paper is the explicit modeling of the state-space model for three-axes stabilized attitude-maneuvering satellite using quaternions, Gibbs vector and MRPs. The second, and the most important contribution, is the optimal factorization of the SDRE technique in an AOCS with nonlinear dynamics for a given set of parameters, initial condition and references for the controller. The models are evaluated for an attitude maneuver called the upside-down in the Launch and Early Orbit Phase (LEOP). In LEOP, the AOCS must dump the residual angular velocity and point the satellite solar panels towards the Sun. The results show that the MRPs provide better performance in the set of scenarios considered.

This paper is organized as follows. In Section II, the problem is described. In Section III, the satellite is modelled, and the basic equations are shared. In Section IV, the state-space models are presented and the SDRE controller is designed. In Section V, the controllability of such models is explored as well as parametrization are evaluated. In Section VI, we investigate the MRPs. In Section VII, some simulations are performed using the optimal factorization. In Section VIII, the conclusions are presented.

II. SDRE METHOD

The SDRE technique entails factorization (that is, parametrization) of the nonlinear dynamics into the state vector and the product of a matrix-valued function that depends on the state itself. In doing so, SDRE brings the nonlinear system to a (not unique) linear structure having SDC matrices given by (1).

$$\dot{\vec{x}} = A(\vec{x})\vec{x} + B(\vec{x})\vec{u}$$
$$\vec{y} = C\vec{x} \quad (1)$$

where x is the state vector and u is the control vector. Notice that the SDC form has the same structure as a linear system, but with the system matrices, A and B, being functions of the state vector. The non uniqueness of the SDC matrices creates extra degrees of freedom, which can be used to enhance controller performance, however, it poses challenges since not all SDC matrices fulfill the SDRE requirements, e.g., the pair (A,B) must be pointwise stabilizable.

The system (1) is subject of the cost functional as in (2).

$$J(\vec{x}_0,\vec{u}) = \frac{1}{2}\int_0^\infty (\vec{x}^T Q(\vec{x})\vec{x} + \vec{u}^T R(\vec{x})\vec{u})dt \quad (2)$$

where Q and R are the state-dependent weighting matrices. In order to ensure local stability, Q is required to be positive semi-definite for all x and R is required to be positive for all x [10]. The SDRE controller linearizes the plant about the current operating point and creates constant state space matrices so that the LQR method can be used. This process is repeated in all samplings steps, resulting in a pointwise linear model from a non-linear model, so that an ARE is solved and a control law is computed also in each step. Therefore, according to LQR theory and (1) and (2), the state-feedback control law in each sampling step is u=-$K(x)x$ and the state-dependent gain K is obtained by (3) [2].

$$K(\vec{x}) = R^{-1}(\vec{x})B^T(\vec{x})P(\vec{x}) \quad (3)$$

where P(x) is the unique, symmetric, positive-definite solution of the algebraic state-dependent Riccati equation (SDRE) given by (4).

$$P(\vec{x})A(\vec{x}) + A^T(\vec{x})P(\vec{x}) - P(\vec{x})B(\vec{x})R^{-1}(\vec{x})B^T(\vec{x})P(\vec{x}) + Q(\vec{x}) = 0 \quad (4)$$

III. SATTELITE EQUATIOS OF MOTIONS

The satellite model is designed based on a typical mission developed by INPE, in which the AOCS must stabilize the satellite in three-axis so that the optical payload can point to the desired target. Therefore, the satellite model is defined to be a three-axis stabilized, attitude-maneuvering satellite, a zero-bias-momentum system. A major control requirement is to remove the unwanted accumulated angular momentum, which would drive the satellite pointing away from the desired target. Thus, an active control system is needed to dump the residual body angular velocity that is created by perturbation torques from the space environment [6][13].

The satellite model available, which is based on the Amazonia-1 [16], uses reaction wheels (momentum exchange actuators) to provide fine attitude control and to maneuver the satellite [13]. The simulator models have two types of sensors to compute and propagate the attitude: (1) a set of Sun sensors, and (2) a gyro, which provide all the necessary information for the LEOP attitude maneuver to acquire the Sun pointing.

The satellite attitude is represented by means of quaternions in the model. Hence, the dynamic equation of

the quaternion Q that rotates the ECI reference frame into alignment with the satellite body reference frame is as in (5) [6].

$$\dot{Q} = \frac{1}{2}\Omega(\vec{\omega})Q = \frac{1}{2}\Xi(Q)\vec{\omega}$$

$$\Omega(\vec{\omega}) \triangleq \begin{bmatrix} 0 & \omega_3 & -\omega_2 & \omega_1 \\ -\omega_3 & 0 & \omega_1 & \omega_2 \\ \omega_2 & -\omega_1 & 0 & \omega_3 \\ -\omega_1 & -\omega_2 & -\omega_3 & 0 \end{bmatrix}$$

$$\Xi(Q) \triangleq \begin{bmatrix} q_4 & -q_3 & q_2 \\ q_3 & q_4 & -q_1 \\ -q_2 & q_1 & q_4 \\ -q_1 & -q_2 & -q_3 \end{bmatrix}, \quad (5)$$

Notice that (5) can be written as in (6), using the vector g (Gibbs vector or Rodrigues parameter) as $Q = [g^T | q_4]$ [6].

$$\dot{Q} = -\frac{1}{2}\begin{bmatrix} \omega^{\times} \\ \omega^T \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} + \frac{1}{2}q_4\begin{bmatrix} I_{3\times3} \\ 0 \end{bmatrix}\vec{\omega} \quad (6)$$

Notice that (5) can be written as in (7) [6], using the vector p as $p = \frac{g}{1+q_4}$.

$$\dot{p} = \frac{1}{2}[\frac{1}{2}(1 - p^T p)I_{3\times3} + p^{\times} + pp^T I_{3\times3}]\omega \quad (7)$$

The model of the rotational dynamics of the satellite is based on the Euler-Newton formulation and considers that the satellite has a set of 3 reaction wheels, aligned with its principal axes of inertia. One can define the inertia moment of the satellite coupled with the 3 reaction wheels by (8).

$$I_b = \vec{I} - \sum_{n=1}^{3} I_{n,s} a_n a_n^T \quad (8)$$

where Ib, I and In,s are the total, the satellite and the inertia moment of the reaction wheels in their symmetry axis an.

Assuming that there is no net external torque and using (8), the rotational dynamics of the satellite is given by (9).

$$\dot{\vec{\omega}} = (-I_b^{-1}\omega^{\times} I_b + I_b^{-1}(\sum_{n=1}^{3} h_{w,n} a_n)^{\times})\vec{\omega} - I_b^{-1}\sum_{n=1}^{3} g_n a_n \quad (9)$$

where, ω is the angular velocity, gn is the torque generated by the n reaction wheel and hw,n is the angular momentum of the n reaction wheel about its center of mass.

IV. DESIGN OF THE SDRE CONTROLLER

The SDRE controller has to deal with two dynamics: (a) the attitude described by unit quaternions Q and (b) the angular velocity, ω, of the satellite. In (a), the attitude must be stabilized and must follow the Sun according to a given Sun vector in the satellite and in (b), the angular velocity read by the gyroscope must be as close as possible to 0. The state and the control vectors are defined by (10).

$$\begin{bmatrix} \vec{x_1} \\ \vec{x_2} \end{bmatrix} = \begin{bmatrix} Q \\ \vec{\omega} \end{bmatrix}$$
$$\begin{bmatrix} \vec{u_1} \end{bmatrix} = \begin{bmatrix} \vec{T_c} = \sum_{n=1}^{3} g_n \vec{a_n} \end{bmatrix} \quad (10)$$

Taking into account the control vector defined in (10), the state space model can be defined using (5), (9) and it is given by (11) and (12).

$$\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} = \begin{bmatrix} Q \\ \omega \end{bmatrix}$$

$$\begin{bmatrix} \dot{x_0} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\Omega & 0 \\ 0 & -I_b^{-1}\omega^{\times} I_b + I_b^{-1}(\sum_{n=1}^{3} h_{w,n} a_n)^{\times} \end{bmatrix}\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -I_b^{-1} \end{bmatrix}\begin{bmatrix} u_1 \end{bmatrix}$$
$$\begin{bmatrix} y \end{bmatrix} = 1\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} \quad (11)$$

$$\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} = \begin{bmatrix} Q \\ \omega \end{bmatrix}$$

$$\begin{bmatrix} \dot{x_0} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2}\Xi \\ 0 & -I_b^{-1}\omega^{\times} I_b + I_b^{-1}(\sum_{n=1}^{3} h_{w,n} a_n)^{\times} \end{bmatrix}\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -I_b^{-1} \end{bmatrix}\begin{bmatrix} u_1 \end{bmatrix}$$
$$\begin{bmatrix} y \end{bmatrix} = 1\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} \quad (12)$$

However, such state-space models, in (11) and (12), do not fulfill the SDRE requirements, in particular, the pair (A,B) is not pointwise stabilizable. Therefore, it is impossible to use such models with SDRE technique.

An alternative option for the definition of the state-space model is to use (6), which leads to (13).

$$\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} = \begin{bmatrix} Q \\ \omega \end{bmatrix}$$

$$\begin{bmatrix} \dot{x_0} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\begin{bmatrix} \omega^{\times} \\ \omega^T \end{bmatrix} & 0 & \begin{bmatrix} \frac{1}{2}q_4 I_{3\times3} \\ 0 \end{bmatrix} \\ 0 & 0 & -I_b^{-1}\omega^{\times} I_b + I_b^{-1}(\sum_{n=1}^{3} h_{w,n} a_n)^{\times} \end{bmatrix}\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -I_b^{-1} \end{bmatrix}\begin{bmatrix} u_1 \end{bmatrix}$$
$$\begin{bmatrix} y \end{bmatrix} = 1\begin{bmatrix} x_0 \\ x_2 \end{bmatrix} \quad (13)$$

(13) has been shown to satisfy SDRE conditions, moreover, in it only A is a function of the state vector, consequently, that is A (x).

 ISBN: 978-1-61208-694-1

Another alternative option for the definition of the state-space model is to use (7), which leads to (14).

$$\begin{bmatrix} x_3 \\ x_2 \end{bmatrix} = \begin{bmatrix} p \\ \omega \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_3 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{4}(1-p^T p)I_{3\times 3} + \frac{1}{2}p^{\times} + \frac{1}{2}(pp^T)I_{3\times 3} \\ 0 & -I_b^{-1}\omega^{\times} I_b + I_b^{-1}(\sum_{n=1}^{3} h_{w,n} a_n)^{\times} \end{bmatrix} \begin{bmatrix} x_3 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -I_b^{-1} \end{bmatrix} \begin{bmatrix} u_1 \end{bmatrix}$$

$$\begin{bmatrix} y \end{bmatrix} = I \begin{bmatrix} x_3 \\ x_2 \end{bmatrix} \quad (14)$$

V. INVESTIGATION OF THE CONTROLABILITY

According to [2], an effective approach for selecting the optimal state-space model for the SDRE is to attempt to maximize the pointwise stabilizability of the possible models, since pointwise control effort can be directly linked to controllability. The controllability criterion requires the value of determinant of the controllability matrix to be different from zero, therefore, a graphical comparison of the absolute value of the determinant of controllability matrix can be used to reveal when pointwise controllability is maximized. For multi-input systems, as the one studied in the present paper, the controllability matrix is not square, and then the controllability matrix multiplied by its transpose is used to evaluate the determinant.

TABLE 1. SATELLITE DATA.

Satellite characteristics	
Inertia tensor (kg.m^2)	$\begin{bmatrix} 310.0 & 1.11 & 1.01 \\ 1.11 & 360.0 & -0.35 \\ 1.01 & -0.35 & 530.7 \end{bmatrix}$
Reaction Wheels characteristics	
Inertia	0.01911
inertia tensor of 3 reaction wheels (kg.m^2)	$\begin{bmatrix} 0.01911 & 0 & 0 \\ 0 & 0.01911 & 0 \\ 0 & 0 & 0.01911 \end{bmatrix}$
Maximum torque (N.m)	0.075
Maximum angular velocity (RPM)	6000
Initial Conditions	
Attitude (degrees, XYZ)	$[0\ 0\ 180]^T$
Angular velocity (radians/second, XYZ)	$[0\ 0\ 0.024]^T$
References for the controller	
Solar vector in the body (XYZ)	$[1\ 0\ 0]^T$
Angular velocity (radians/second, XYZ)	$[0\ 0\ 0]^T$

Numerical simulations were performed to determine which of the (13) or (14) maximizes the controllability of the system for the satellite used which data are shown in Table 1. Figure 1 shows the controllability of the state-space model (SSM) defined by (13) using quaternions and the Gibbs vector, whereas Figure 2 shows the controllability of the same state-space model defined by (14).

From Figures 1 and 2, MRPs maximize the controllability through the simulation. Once MRPs provide better controllability, the next step is to further parametrize such state-model using different SDCs, which is explored in the next section below.

VI. PARAMETRIZATION OF MRP

For multivariable state-space models, as the one studied here, there are two distinct SDC matrices A1(x) and A2(x), once there is an infinite number of SDC parametrizations. Such infinite parametrizations can be constructed using (15).

$$A(x,\alpha) = \alpha A_1(x) + (1-\alpha)A_2(x) \quad (15)$$

where α is a real number and $0 \le \alpha \le 1$.

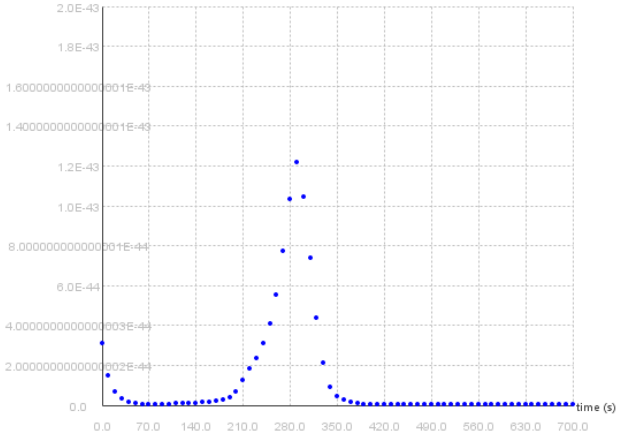


Figure 1. Controllability using quaternions and Gibbs vector.

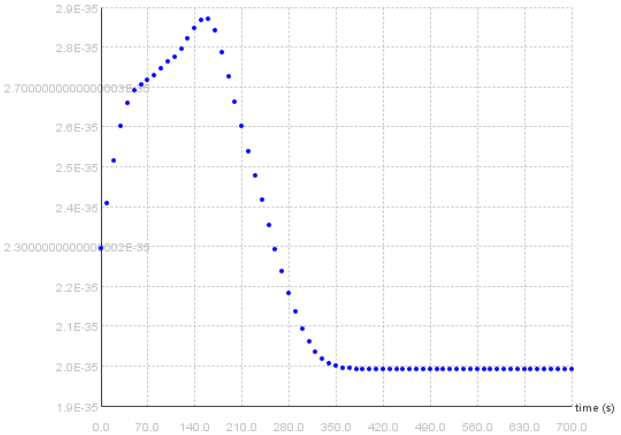


Figure 2. Controllability using MRPs.

Focusing on MRPs, which provide better controllability, and taking into account the rule of thumb that the selection of the state-dependent factorization should place a nonzero entry in the {i,j}-element of the A(x,α) matrix if the ith state derivative depends on the jth state, the state-space model is defined in (16), which is based on algebraic manipulations of (7).

$$\begin{bmatrix} x_3 \\ x_2 \end{bmatrix} = \begin{bmatrix} p \\ \omega \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_3 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\omega^{\times} + \frac{1}{2}(\omega^T p)I_{3\times 3} & \frac{1}{4}(1-p^T p)I_{3\times 3} \\ 0 & -I_b^{-1}\omega^{\times} I_b + I_b^{-1}(\sum_{n=1}^{3} h_{w,n} a_n)^{\times} \end{bmatrix} \begin{bmatrix} x_3 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -I_b^{-1} \end{bmatrix} \begin{bmatrix} u_1 \end{bmatrix}$$

$$\begin{bmatrix} y \end{bmatrix} = I \begin{bmatrix} x_3 \\ x_2 \end{bmatrix} \quad (16)$$

Figure 3 shows the controllability of the state-space model defined by (16) based on MRPs, which is better than the one shown in Figure 1 so MRPs still have better controllability than quaternions and the Gibbs vector. Nevertheless, this second option based on MRPs (16) has worse controllability than the first one defined by (14).

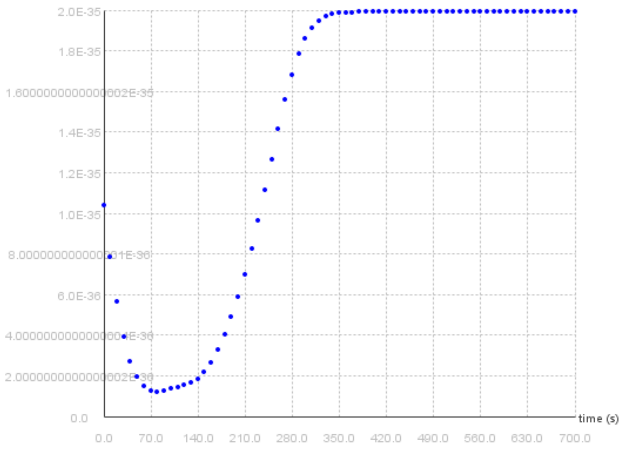


Figure 3. Controllability of MRPs of (16).

The two distinct parametrizations of (14) and (16) for the same system suggest the possibility of using (15) to evaluate a new combination of these two provides better controllability, that new parametrizations using A is given by (17).

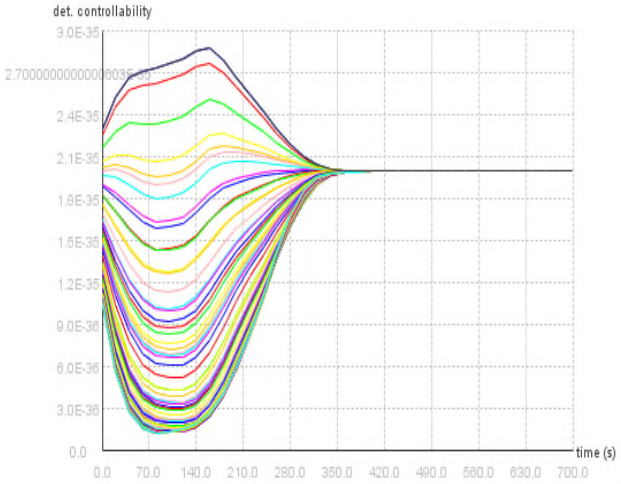
$$A_{1_{x_3}}(x_3) = \begin{bmatrix} -\frac{1}{2}\omega^{\times} + \frac{1}{2}(\omega^T p)I_{3\times 3} & \frac{1}{4}(1-p^T p)I_{3\times 3} \end{bmatrix}$$

$$A_{2_{x_3}}(x_3) = \begin{bmatrix} 0 & \frac{1}{4}(1-p^T p)I_{3\times 3} + \frac{1}{2}p^{\times} + \frac{1}{2}(pp^T)I_{3\times 3} \end{bmatrix} \quad (17)$$

Applying the new parametrizations (17) into (15) and to compare the performance of resulting parametrization with different α, a simulation test was conducted with the full Monte Carlo perturbation model, in which α was randomly selected in the interval $0 \leq \alpha \leq 1$. The Monte Carlo model ran 90 times. Each time, the simulation used a different α with the data of Table 1. Figure 4 shows the resulting controllability of each run. It is possible to conclude that the parametrization defined by (14) is optimal since the controllability is the highest through the entire simulation. Therefore, the state-space model defined by (14), resulting in the controllability shown in Figure 2, is the optimal factorization to design the SDRE controller with nonlinear dynamics. Such conclusion is based on the characteristics of the satellite, the initial conditions and the references for the controller in Table 1. So, this is not valid for the general case and it is not valid for a different initial condition. Moreover, as (14) is defined using MRPs, which have singularity for 360o, it is neither unique nor global, whereas (13) based on quaternions and Gibbs vector is global but not unique.

VII. SIMULATION USING OPTIMAL FACTORIZATION

The satellite control using the SDRE controller designed with the optimal factorization is shown in Figures 5 and 6.



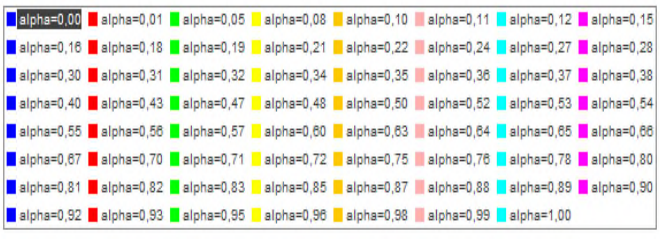


Figure 4. Controllability of MRPs using state-space model (14) and (16).

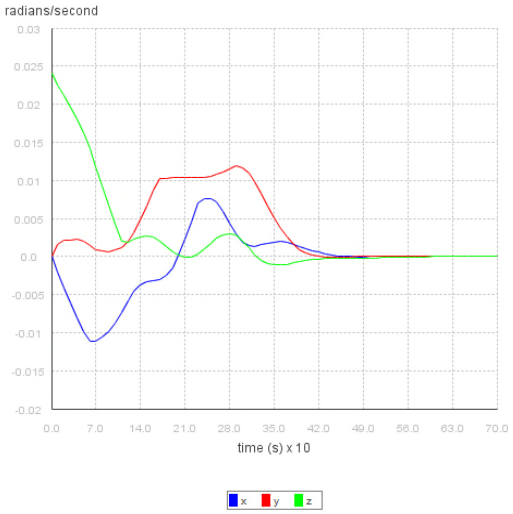


Figure 5. Control of satellite angular velocity with the optimal factorization.

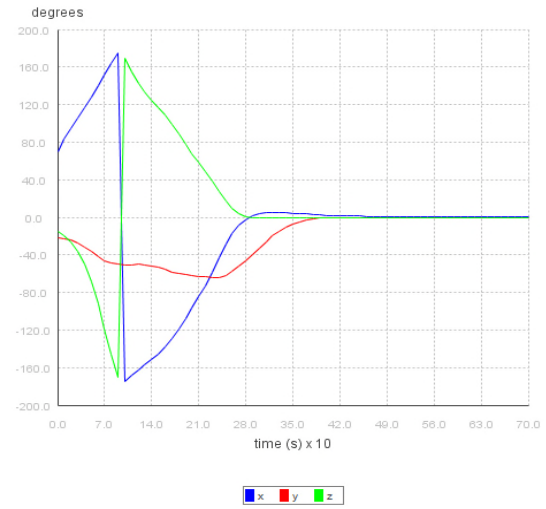


Figure 6. Control of satellite attitude with the optimal factorization.

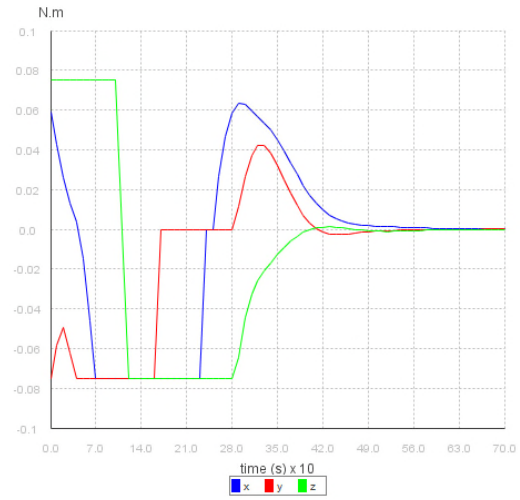


Figure 7. Reaction wheels torque with the optimal factorization.

Focusing on the actuators, Figure 7 shows that the reaction wheels torque is stabilized at about 420s.

VIII. CONCLUSION

This is a contribution for the optimal arrangement of the SDC for a three-axis stabilized satellite model. The results shown that different SDCs can produce extremely different results ranging from non-applicability of the SDRE technique to huge differences in the controllability and, consequently, in the performance and robustness of the system. Unfortunately, the optimal factorization found is neither valid for the general case nor even for a different initial condition due to the underlining nonlinear dynamics. However, the procedure applied can provide guidance for engineers. A secondary contribution is the kinematical part of the state-space models in (13), (14) and (16), since they can be used in any system that exhibits rotational motion, e.g., airplanes.

The simulations shown in the figures were performed using a portable simulator developed at INPE [5,15]; such simulator has capabilities for the unloading of the angular momentum of the reaction wheels (based on a magnetometer and a set of magnetorquers) not explored in the current paper.

One aspect discussed in this paper was whether the SDC factorization of SDRE technique in AOCS design can yield gains in the missions developed by INPE. Since performance in the LEOP is critical to the success of a mission and the simulation results show that the performance and robustness of SDRE controllers can be enhanced by optimal factorizations (in particular, with kinematics based on MRPs), then we can say that SDRE can yield gains in the missions developed by INPE. Nonetheless, its implementation requires more computing resources and tends to exhibit difficulties for verification. Therefore, it is too early to draw a definitive conclusion about the applicability of SDRE in missions at INPE. However, we can conclude that once SDRE technique is used, the optimal factorization of SDC is of utmost importance for performance and robustness of nonlinear systems controlled by such technique.

REFERENCES

[1] T. Cimen, “State-Dependent Riccati Equation (SDRE) Control: A Survey”, IFAC Proceedings Volumes, Vol. 17, pp. 3761-3775, 2008.

[2] T. Cimen, “Systematic and effective design of nonlinear feedback controllers via the state-dependent Riccati equation (SDRE) method”, Annual Reviews in Control, Vol. 34, pp. 32-51, 2010.

[3] J. R. Cloutier, C. N. D’Souza, and C. P. Mracek, “Nonlinear regulation and nonlinear H-infinity control via the state-dependent Riccati equation technique”, Proceedings of the First International Conference on Nonlinear Problems in Aviation and Aerospace, pp. 117-141, 1996.

[4] R. G. Gonzales and L.C. G. Souza, ”Application of the SDRE Method to Design an Attitude Control System Simulator”,

 ISBN: 978-1-61208-694-1

Advances in the Astronautical Sciences, Vol. 134, p. 2251-2258, 2009.

[5] Hipparchus.org, “Hypparchus 1.3-SNAPSHOT”, https://github.com/Hipparchus-Math/hipparchus, Access date: March.2019.

[6] P. C. Hughes, “Spacecraft Attitude Dynamics”, NY, 1986.

[7] A. Jagat and A. J. Sinclair, “Nonlinear Control for Spacecraft Pursuit-Evasion Game Using SDRE Method”, IEEE Transactions on Aerospace and Electronic Systems, Vol. 53, pp. 3032-3042, 2017.

[8] G. D. Mauro, M. Schlotterer, S. Theil, and M. Lavagna, “Nonlinear Control for Proximity Operations Based on Differential Algebra”, Journal of Guidance, Control, and Dynamics, Vol. 38, pp. 2173-2187, 2015.

[9] G. D. Mauro, P. D. Lizia, and M. Lavagna, “Control of Relative motion via State-Dependent Riccati Equation Technique”, Proceedings of the AAS/AIAA Astrodynamics Specialist Conference in Advances in Astronautical Sciences, Vol. 142, pp. 909-928, 2012.

[10] P. K. Menon, T. Lam, L. S. Crawford, and V. H. L. Cheng, “Real-time computational methods for SDRE nonlinear control of missiles”, Proceedings of the 2002 American Control Conference, Vol. 1, pp. 232-237, 2002.

[11] J. D. Pearson, “Approximation Methods in Optimal Control”, Journal of Electronics and Control, Vol. 13, pp. 453-469, 1962.

[12] C. Pukdeboon, “Optimal Sliding Mode Controllers for Attitude Stabilization of Flexible Spacecraft,” Mathematical Problems in Engineering, Vol. 2011, Article ID 863092, 20 pages, 2011.

[13] M. J. Sidi, “Spacecraft Dynamics and Control - A Practical Engineering Approach”, Cambridge Univ. Press, 2006.

[14] D. T. Stansbery and J. R. Cloutier, "Position and attitude control of a spacecraft using the state-dependent Riccati equation technique," Proceedings of the 2000 American Control Conference, Vol. 3., pp. 1867-1871, 2000.

[15] A. G. Romero and L. C. G. Souza, “Satellite Controller System Based on Reaction Wheels Using the State-Dependent Riccati Equation (SDRE) on Java”, Proceedings of the 10th International Conference on Rotor Dynamics – IFToMM, Vol. 1, pp. 547-561, 2018.

[16] Inpe.br, “Missão Amazônia 1”, http://www3.inpe.br/amazonia-1/, Access date: March, 2019.

 ISBN: 978-1-61208-694-1

Interference-Free Store-and-Forward Communication in Low Earth Orbit Satellite Systems

Andreas Freimann*, Timon Petermann† and Klaus Schilling‡
University of Würzburg, Germany
Email: *freimann@informatik.uni-wuerzburg.de
†timon.petermann@stud-mail.uni-wuerzburg.de
‡schi@informatik.uni-wuerzburg.de

***Abstract*—Multi-satellite systems in low Earth orbits based on nanosatellites can be used for Earth observation and global store-and-forward communication. The intersatellite links in these systems need to be coordinated in a way that avoids interference, and optimises throughput, delays and energy consumption. Existing contact plan design approaches do not prevent packet loss due to interference of nodes, and the use of additional Medium Access Control (MAC) schemes such as Carrier-sense multiple access with collision avoidance (CSMA/CA) is inefficient in networks with high propagation time and dynamic topology. In this paper, a contact plan based design approach is presented that solves the MAC problem by predicting additive interference and scheduling interference-free inter-satellite links in advance. The presented approach is part of a Delay Tolerant Network (DTN) routing protocol for space-terrestrial networks. An improvement of the algorithm is also presented to reduce computational effort. Performance was evaluated for different orbit configurations using a satellite network simulation framework based on the Objective Modular Network Testbed in C++ (OMNeT++).**

Keywords–Satellite Communication; Delay Tolerant Network; Medium Access Control; Satellite Formations.

I. INTRODUCTION

The most promising application scenarios for multi-satellite systems based on nanosatellites are Earth Observation (EO) and Machine-to-Machine (M2M) communication. Nanosatellite based systems for Earth observation are already in use, such as the constellation of Planet [1] with its Dove satellites. In the future, satellite formations will be launched that capture 3D information, such as CloudCT [2], a formation of ten satellites build by the Center for Telematics (ZfT) Würzburg. In this mission, 3D information of clouds will be acquired by a satellite formation flying in a dense orbit configuration. For satellite based communication systems, spatially distributed orbit configurations are used due to their high spatial and temporal coverage of the Earth's surface. Accordingly, a globally distributed satellite system and a dense satellite formation have been used for evaluation of the presented approach.

Both M2M and EO satellite systems could take advantage of the development of efficient communication protocols by increasing throughput, faster availability of data and more efficient use of resources. The goal of the presented approach is to take advantage of the predictability of the dynamic topology of satellite networks for the design of routing algorithms. In the next section, the basic principle of the presented approach will be described, before, in Section III, a contact plan design approach, called Interference-Free Contact Plan Design (IFCPD) is presented. The performance of this approach is evaluated in Section IV, followed by a conclusion in Section V.

II. DTN ROUTING

Communication in low Earth orbit satellite systems can benefit from Delay/Disruption Tolerant Network (DTN) protocols to cope with intermittent and unreliable links. Routing in such a network can be difficult, but routing algorithms can make use of the knowledge about a system. The motion of satellites is highly predictable. Instead of a topology discovery mechanism, contact plan based routing utilises predefined contact plans of the network. Therefore nodes can make routing decisions based on future contacts within the network. So far no Multiple Access Control (MAC) scheme is available that makes use of the available contact plans. Therefore standard MAC schemes need to be used to prevent packet loss, such as CSMA/CA, which is inefficient in Low Earth Orbit (LEO) networks as shown in [3]. As a basis for the presented approach, a simplified general perturbations (SGP4) propagator [4] was integrated into an OMNeT++ based simulation framework to predict satellite positions. By propagating satellite positions, the network topology and the node distances can be calculated with sufficient accuracy for days in advance. The predicted topology is stored in a contact plan that consists of multiple contact entries. A contact is defined by its start time, end time, source node, sink node and a data rate. The contact plan is converted to a space time graph for routing purposes. The vertices of a space time graph are network nodes. Edges are either the contacts from the contact plan or edges between copies of the same network nodes representing storage of data on these nodes. In this graph, the shortest routes for all pairs of nodes are computed and stored in routing tables. For more details on this approach please refer to [3]. The presented contact plan design approach is compatible with Contact Graph Routing (CGR) [5], a contact plan based routing approach for the Bundle Protocol [6].

III. INTERFERENCE-FREE CONTACT PLAN DESIGN

The satellites are assumed to be equipped with isotropic antennas and typical Ultra High Frequency (UHF) transceivers with a transmit power of $0.5W$. Transmissions are assumed to be successful if the received signal-to-interference-plus-noise ratio (SINR) is at least $5.82dB$.

To create a contact plan for a specific orbit configuration, a check for all pairs of nodes is performed, on whether these pairs are able to communicate with each other. For each of these contacts, potential interfering nodes are identified by iterating over all transmitters that exist in the simulation scenario. In a system with n nodes, there are $n-2$ nodes that

 ISBN: 978-1-61208-694-1

could generally cause interference on a link, so $n-2$ interference checks have to be performed to identify which nodes would cause interference on a link at a specific point in time. Furthermore constructive interference of signals produced by two or more nodes could cause packet loss, which is also called additive interference. To identify all sets of interfering links, $2^{(n-2)}-1$ checks need to be performed in the worst case. In scenarios with a high number of network nodes and long prediction intervals additive interference checks require a very high computational effort. However, the number of checks can be reduced by more than half for each node that can be excluded from the list of potential interfering nodes. To improve performance, all nodes are stored in a spatial data structure (k-d tree) so that a radius search can be applied with the radius equal to the line-of-sight distance. Thereby all nodes that are not separated by Earth can be picked for interference evaluations. The radius search is also applied during contact plan creation by using the communication range as radius. Using a contact and an interference plan, maximal sets of non-interfering contacts can be calculated. These independent sets can be determined by the following algorithm:

Initialization: Create a bin for each contact within a contact plan time interval and a list of the succeeding contacts.

1) For all bins generated in the previous iteration, iterate through the contacts of the corresponding list, create a copy of the bin, add the selected contact and a list of the succeeding contacts.
2) For all bins generated in the previous step, check if the last contact in the bin interferes with any subset of the contacts in the bin. Delete the corresponding bin if interference was detected.
3) If all lists are empty stop here, otherwise continue with step 1.

Finally, check for redundant bins and remove them.

IV. Evaluation

Evaluations have been performed in two mission scenarios and compared by runtime and packet delivery ratios.

A. Runtime comparison

To evaluate the performance improvement of the radius search, a 40 minute contact and interference plan with additive interference checks for a $45° : 18 : 6 : 0$ walker constellation at an altitude of 698km has been created on a standard PC. The calculations have been performed for a period of 40 minutes. The runtime with radius search was 0.56s, compared to 342.86s without radius search. If just a contact plan was generated for one day the runtime could be reduced from 4.47s to 1.16s. This proves that using a spatial data structure with an efficient radius search increases the performance of contact and interference plan calculations significantly. This improvement is mainly important to be able to consider additive interference in larger satellite systems.

B. Packet loss

To evaluate the packet delivery ratio of the IFCPD algorithm, a 10 satellite string-of-pearls formation comparable to the CloudCT mission was analysed. The simulated altitude is 600km and the node separation is 100km. In the simulated scenario, each satellite tries to transmit data to its successor during a 10,000s simulation period. The satellites transmit one packet every 10s simultaneously to focus on the analysis of situations where packet loss would occur due to interference.

TABLE I. Packet delivery ratios of the evaluated algorithms

Algorithm	Packet Delivery Ratio
FCP	0%
IACPD	97.6%
IFCPD	100%

The IFCPD approach was compared to the previous version, called Interference-Aware Contact Plan Design (IACPD), presented in [3] and to the basic CGR [5] approach with a Full Contact Plan (FCP) that includes all possible communication opportunities. IACPD does only consider interference caused by single satellites, while the use of a FCP does not avoid interference at all.

Table I shows the summarised evaluation results. In case a full contact plan is used, 100% of the transmitted packets are lost due to interference. IFCPD is able to deliver all packets, while IACPD produces a packet loss of 2.4%.

V. Conclusion

The presented contact plan design approach IFCPD was able to prevent packet loss in simulated wireless satellite networks and seems to be a suitable and efficient approach for LEO networks. The presented improvement of IFCPD based on radius search significantly reduces the computational effort. Nevertheless, performance of the algorithm for the calculation of independent sets has to be improved to enable calculation of plans for scenarios with a higher number of nodes and longer prediction intervals. Furthermore, since system models are used to predict future states of a system, the performance of the presented approach will depend on the quality of these models. Therefore, we will further investigate the robustness of the approach in case of model errors.

Acknowledgement

This work was supported by the German Research Foundation (DFG) under grand number SCHI 327/29-1.

References

[1] C. Boshuizen, J. Mason, P. Klupar, and S. Spanhake, "Results from the Planet Labs Flock Constellation," in *28th Annual AIAA/USUConference on Small Satellites*, 2014.

[2] ERC Synergy Grant "CloudCT" awarded. Last accessed 18 March 2019. [Online]. Available: http://www7.informatik.uni-wuerzburg.de/news/single-news/news/erc-synergy-grant-cloudct-awarded

[3] T. Thiel, A. Freimann, and K. Schilling, "Contact plan based routing in distributed nanosatellite systems," in *69th International Astronautical Congress (IAC)*, 2018.

[4] D. A. Vallado and P. Crawford, "SGP4 orbit determination," in *Proceedings of AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2008, pp. 18–21.

[5] G. Araniti et al., "Contact graph routing in DTN space networks: overview, enhancements and performance," *Communications Magazine, IEEE*, vol. 53, no. 3, pp. 38–46, 2015.

[6] S. Burleigh and K. Scott. (2007) Bundle protocol specification. Last accessed 10 December 2018. [Online]. Available: http://tools.ietf.org/html/rfc5050

 ISBN: 978-1-61208-694-1

High Sensitivity Inter-Satellite Optical Communications Using LEDs and Single Photon Receivers

Alexander D. Griffiths, Johannes Herrnsdorf, Michael J. Strain and Martin D. Dawson

Institute of Photonics, Dept. of Physics
University of Strathclyde
Glasgow, UK
Email: alex.griffiths@strath.ac.uk, johannes.herrnsdorf@strath.ac.uk,
michael.strain@stratch.ac.uk, m.dawson@strath.ac.uk

***Abstract*—High data rate communication links are difficult to perform with radio frequency technology under the strict Size, Mass and Power (SMaP) constraints of small satellites. Transmitting data using the optical spectrum can permit higher data rates with more power efficient systems. Single photon avalanche diodes (SPADs) provide receivers with exceptional sensitivity levels, while Light-Emitting Diodes (LEDs) bonded to electronic driver chips provide digitally controllable optical transmitters capable of high bandwidth spatial and temporal modulation. Data rates of 100 Mb/s have been demonstrated at a sensitivity of -55.2 dBm, and ray-tracing simulations indicate ranges in excess of 1 km are feasible with simple optical systems. Additionally, the divergent nature of LEDs can provide a level of spatial coverage, relaxing pointing and alignment requirements. The low electrical power requirements and compact, semiconductor nature of these devices may therefore bring high data rate, high sensitivity communications to small satellite platforms.**

Keywords–Light-emitting diodes; Single-photon avalanche diodes; Optical communications.

I. INTRODUCTION

In recent years, the development and deployment of small satellites has increased rapidly, with the development of the CubeSat standard greatly increasing academic and commercial access to space [1]. The future deployment of constellations, clusters or networks of small satellites necessitates high-speed inter-satellite data connections operating under strict Size, Mass and Power (SMaP) budgets. Communicating using the optical spectrum is an attractive alternative to radio frequencies due to the potential for high data rate systems with lower SMaP requirements [2]. Optical Inter-Satellite Links (ISLs) for large satellites achieve long ranges and Gb/s data rates using phase coherent laser communication methods, but are complex, costly and arguably too large and power hungry for small satellites [3].

Research in optical communications for CubeSats has been focussed on satellite to ground links using Laser Diodes (LDs) and fibre amplifiers as transmitters, however, similar systems have been proposed for ISLs [4] and a low-earth orbit (LEO) relay system [5]. The SMaP requirements for these communication systems are well within the capability of CubeSats, though they typically occupy at least a full 1U module and require significant battery power. Light-Emitting Diodes (LEDs) are a potential alternative device suitable for ISLs [6], bringing further advantages in SMaP constraints over laser systems, along with reduced complexity, longer lifetimes and lower cost. Gallium Nitride (GaN) micro-LEDs, LEDs with dimensions below $100\,\mu\mathrm{m}$, show high modulation bandwidths and are an emerging technology for mass market applications. Previous research, including from our own group, suggests data rates comparable to the CubeSat LD systems are achievable [7]. Additionally, the angular divergent nature of LED emission may relax the tight pointing requirements found in laser systems.

Here, we present an experimental demonstration of high sensitivity, intensity modulated optical communications based on LEDs and silicon Single-Photon Avalanche Diode (SPAD) arrays. The advantage of such a system is in its simplicity and low SMaP requirements, while still achieving high sensitivities and data rates. To justify that the ranges, data rates, and angular tolerances are useful, ray-tracing simulations are presented.

In Section II, the experimental performance and SMaP characteristics of the communication link are described, followed by the ray-tracing results. Conclusions and further work are presented in Section III.

II. MAIN RESULTS

The optical transmitter used in the experimental work is a single $99 \times 99\ \mu\mathrm{m}^2$ pixel, one of a 16×16 array of GaN micro-LEDs bump bonded to complementary metal-oxide-semiconductor (CMOS) control electronics. This device provides a mm-scale chip containing optical emission elements and driver electronics for independent control of every pixel. Fabrication and characterisation details of similar devices are reported in [8]. The emission wavelength is centred on 450 nm, primarily due to device availability. We note that such devices can be fabricated from the ultraviolet-C (UV-C) band to green using nitride alloys with wavelengths spanning many of the solar Fraunhofer lines, which may offer a low background noise channel [6]. The receiver is a 64×64 array of SPADs, arranged on a $21\,\mu\mathrm{m}$ pitch with a 43% fill factor [9]. The digital photon count signals from the pixels are combined through XOR trees and ripple counters, causing the device to operate as a digital silicon photomultiplier. An analytical discussion of the performance of SPAD arrays for optical communications can be found in Reference [10].

Details of the experimental arrangement and characterisation can be found in [11]. The resulting receiver sensitivity results, -60.5 and -55.2 dBm for 50 and 100 Mb/s, respectively, are shown in Figure 1 in red squares. Also shown are sensitivities achieved by complementary work from [9], where the same receiver was used for On-Off Keying (OOK) with a LD transmitter. The experimental results remain approximately a constant separation in dB from the standard quantum limit, set by the Poissonian nature of photon detection. The linear

 ISBN: 978-1-61208-694-1

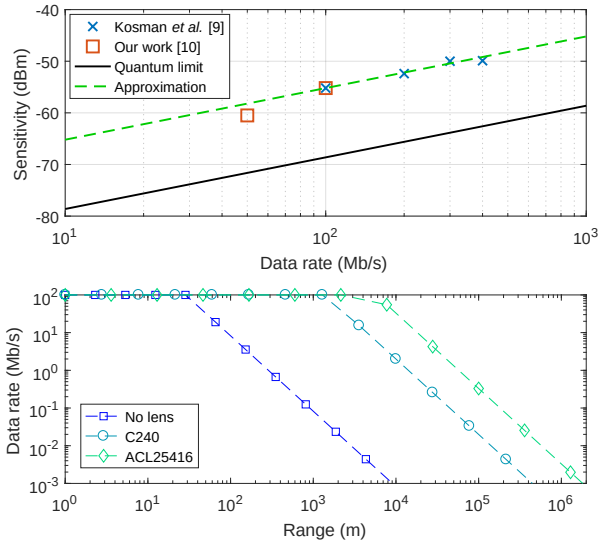


Figure 1. (a) Sensitivity limit for given data rates, experimental results and an approximation. (b) Achievable data rates in a point to point link, based on received intensities and the approximation in (a).

approximation shown in Figure 1 is used in the following simulations to estimate data rates for calculated levels of incident power.

Both the SPAD and LED chips are packaged and mounted on evaluation printed circuit boards (PCBs) where they are controlled and powered through field-programmable gate arrays (FPGAs) (Opal Kelly XEM3010/XEM6310). This results in transmitter and receiver package sizes of 14×18.5 cm^2 and 12.5×20.5 cm^2 respectively. Large parts of these evaluation boards are unnecessary for final applications, so a full transceiver system could potentially be developed on the CubeSat 10×10 cm^2 standard. The electrical power consumption of the full system under operation totals 5.48 W, with the majority consumed by the FPGA control systems. The system has not yet been optimised for power consumption, so this value can be considered an upper limit. Many CubeSats already employ FPGAs as part of their on-board systems, which could be used to produce and process the digital signals required, meaning only the micro-LED and SPAD arrays themselves would consume additional power, estimated to total less than 1 W.

In order to determine a range and coverage performance envelope for the system, ray tracing simulations were performed using Zemax OpticStudio®. The transmitter was specified as a 100×100 µm square, emitting 1 mW with a Lambertian profile typical of LEDs. A 10 cm diameter collection area was assumed at the receiver, likely the largest on-board collection area achievable with a CubeSat system. Three transmitter optical systems were considered: (i) no lens, (ii) collimation with the Thorlabs C240 lens used in the practical experiments, and (iii) collimation with a Thorlabs ACL25416 lens. The simulations were set up to determine intensity maps at varying ranges up to 10^6 m. From the sensitivity approximation in Figure 1(a), achievable data rates can then be determined. The point-to-point results are shown in Figure 1(b). The intensity falls off following the inverse quadratic trend expected for the diverging beam, resulting in a reduction of achievable data rate. Note that the data rate has been upper-end clipped at 100 Mb/s as this is the limit achievable with the micro-LED transmitter operated in this manner. The results suggest the communication link can provide useful data rates over ranges up to 100's of kilometres. Range and data rate requirements will vary strongly with different satellite application scenarios, however these results are in line with other targets in the literature [4]–[6].

The collimation of light from the LED is still divergent, causing losses and limiting range. Importantly, this divergent light can be thought of as providing a level of coverage over an angular region, reducing the pointing accuracy required of the satellites. With the simulated intensity data, the achievable range for a given data rate at an angle from the transmitter can be determined. The results are shown in Figure 2 as angular plots. Note that the radial axis is logarithmic, and in the case of Figure 2(b) and Figure 2(c), the angular axis has been expanded in order to observe the profiles. With no lens, the micro-LED provides a very wide degree of angular coverage due to the Lambertian emission profile. For example, a 100 Mb/s link can be maintained at almost 60 degrees over a 10 m range. At 10 kb/s, the same angular coverage is maintained at approximately 1 km. The coverages provided by the collimated beams in Figure 2(b) and Figure 2(c) are much lower, at roughly 1 and 0.4 degrees for the C240 and ACL25416 respectively.

In a real deployment of this communication system on a CubeSat platform, additional constraints and factors will influence the performance of the ISL. Additional background light incident on the SPAD receiver increases the signal power required to maintain error free performance, so care would have to be taken to avoid background light sources through spectral and spatial filtering. In extreme cases, such as under direct solar irradiation, the receiver will likely saturate rendering the communication link useless. Secondly, the receiver optics have been simulated as a perfect collection area, a practical system will have physical limitations on achieving this, and some degree of loss will be inevitable. Finally, the performance of these bespoke devices in orbit, with potential for cosmic ray damage, is unknown. Silicon devices have a long history in space applications, and the avalanche nature of SPAD operation may help in protecting them from catastrophic breakdown. GaN devices are less mature for space applications, though the fact that they function in spite of high numbers of crystal dislocations suggests some degree of resilience to defect-causing damage from charged particles.

III. Conclusion

In summary, a simple, low SMaP and high sensitivity communication link based on micro-LEDs and SPADs provides an attractive platform for CubeSat ISLs. The link distance and coverage simulations shown here suggest ranges up to 100s of kilometres are achievable with significantly relaxed pointing requirements compared to laser systems. Future experimental work will initially focus on a practical demonstration of communication at longer ranges (> 100 m) through air. Significant work will be required to build a system suitable for testing in orbit on a CubeSat, primarily in developing an interface with on-board computer systems and adjusting hardware to the correct form factor. Once a final system has been developed, ray-tracing simulations with accurate parameters, such as emitted power, collection aperture and chosen optics, can be performed

 ISBN: 978-1-61208-694-1

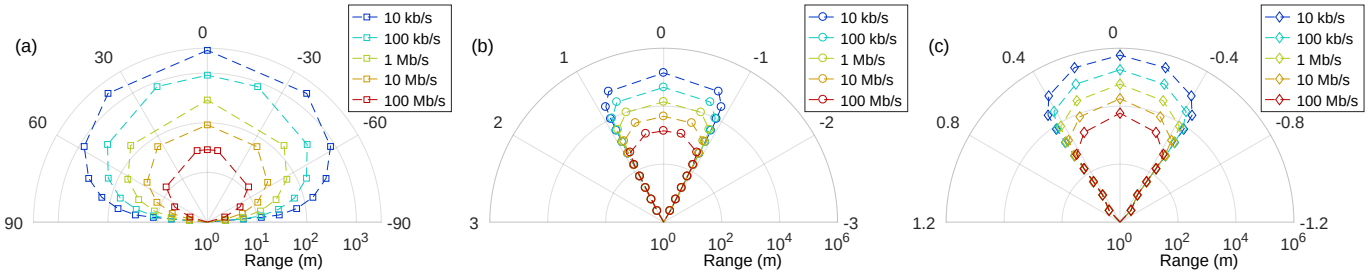


Figure 2. Angular coverage and range performance at data rates from 10 kb/s to 100 Mb/s for (a) no lens, (b) C240 and (c) ACL25416 cases. Note the angular ranges in (b) and (c) have been adjusted to aid visualisation of the narrow beam cases.

to predict performance ahead of an in-orbit demonstration.

The underlying data for this work is available at https://doi.org/10.15129/cb3af47d-3d40-49d6-8cf6-52c13fcf1769

ACKNOWLEDGMENT

This work was funded by the UK Engineering and Physical Sciences Research Council (EPSRC) EP/M01326X/1. We acknowledge Robert Henderson and Oscar Almer from the University of Edinburgh for development of the SPAD array.

REFERENCES

[1] M. N. Sweeting, "Modern Small Satellites - Changing the Economics of Space," Proceedings of the IEEE, vol. 106, no. 3, 2018, pp. 343–361.

[2] H. Kaushal and G. Kaddoum, "Optical Communication in Space: Challenges and Mitigation Techniques," IEEE Communications Surveys and Tutorials, vol. 19, no. 1, 2017, pp. 57–96.

[3] Z. Sodnik, B. Furch, and H. Lutz, "Optical inter-satellite communication," IEEE Journal of Selected Topics in Quantum Electronics, vol. 16, no. 5, 2010, pp. 1051–1057.

[4] R. Morgan and K. Cahoy, "Nanosatellite Lasercom System," AIAA/USU Conference on Small Satellites, 2017, pp. 1–9.

[5] R. Welle et al., "A CubeSat-Based Optical Communication Network for Low Earth Orbit," AIAA/USU Conference on Small Satellites, 2017, pp. 1–9.

[6] D. N. Amanor, W. W. Edmonson, and F. Afghah, "Inter-Satellite Communication System based on Visible Light," IEEE Transactions on Aerospace and Electronic Systems, vol. 54, no. 6, 2018, pp. 2888–2899.

[7] S. Rajbhandari et al., "A review of gallium nitride LEDs for multi-gigabit-per-second visible light data communications," Semiconductor Science and Technology, vol. 32, no. 2, 2017, pp. 1–44.

[8] S. Zhang et al., "1.5 Gbit/s multi-channel visible light communications using CMOS-controlled GaN-based LEDs," Journal of Lightwave Technology, vol. 31, no. 8, 2013, pp. 1211–1216.

[9] J. Kosman et al., "A 500Mb/s -46.1dBm CMOS SPAD receiver for laser diode visible-light communications," IEEE International Solid-State Circuits Conference (in press), 2019, pp. 15–17.

[10] L. Zhang et al., "A comparison of APD and SPAD based receivers for visible light communications," Journal of Lightwave Technology, vol. 36, no. 12, 2018, pp. 2435–2442.

[11] A. D. Griffiths et al., "High-sensitivity free space optical communications using low size, weight and power hardware," ArXiv e-prints, 2019, arXiv:1902.00495 [physics.app-ph].

 ISBN: 978-1-61208-694-1