Improving Critical Communications in Northern Canada

Paul Labbé
National Defence
Defence Research and Development Canada
Ottawa, Canada
email: Paul.Labbe@drdc-rddc.gc.ca / Paul.Labbe@forces.gc.ca

***Abstract*—Evolving activities in the Canadian Arctic drive the need for increased dependable high-data rate communication capabilities with low latencies. This study examines performance and challenges of past and current communication technologies available in this northern region for sensor data, video and voice exchange. Some options for operations in the Canadian Arctic are explored accounting for adverse conditions, such as atmospheric disturbances (both natural and man-made) or adversarial attacks on satellites and terrestrial infrastructure. Potential users include Canadian Armed Forces (CAF), North American Aerospace Defense Command (NORAD), off-grid communities and Public Safety, with respective systems requiring machine-to-machine low latency data sharing. Technologies considered include satellites, microwave relays, fiber optic links, radios such as cellular phones, transceivers in high frequency bands (20-30 Mhz), and particularly Unmanned Aerial System (UAS) gateways.**

Keywords-communications; satellite; UAS; Arctic; latency.

I. INTRODUCTION

This invited paper expands from [1] to examine telecommunications options for operations in Northern Canada accounting for adverse conditions such as atmospheric disturbances (both natural and man-made) as well for adversarial attacks on satellites and terrestrial infrastructure. While most sections of this paper can be highly scientific, some layman analogies have been added throughout to reach a broader audience.

This Defence Research and Development Canada (DRDC) study was initiated to address Canadian Arctic communication challenges expressed in the new Canada's Defense Policy, *Strong, Secure, Engaged* (SSE) [2], which reaffirmed Canada's commitment to effective operations in the Arctic. SSE defines an extended Canadian Air Defense Identification Zone (CADIZ), which includes the entire Canadian Arctic Archipelago (Figure 1), in respect of overall North American Aerospace Defense Command (NORAD) modernization efforts towards an improved North Warning System (NWS) requiring high-throughput low-latency communications.

Some aspects of current Northern Canadian communication systems of the Canadian Armed Forces (CAF), off-grid communities, Search And Rescue (SAR), and Public Safety (PS) are identified. A rise in commercial interest, research and tourism in this zone brings increased safety and security demands to address SAR and natural or man-made disasters to which Canada must be ready to respond.

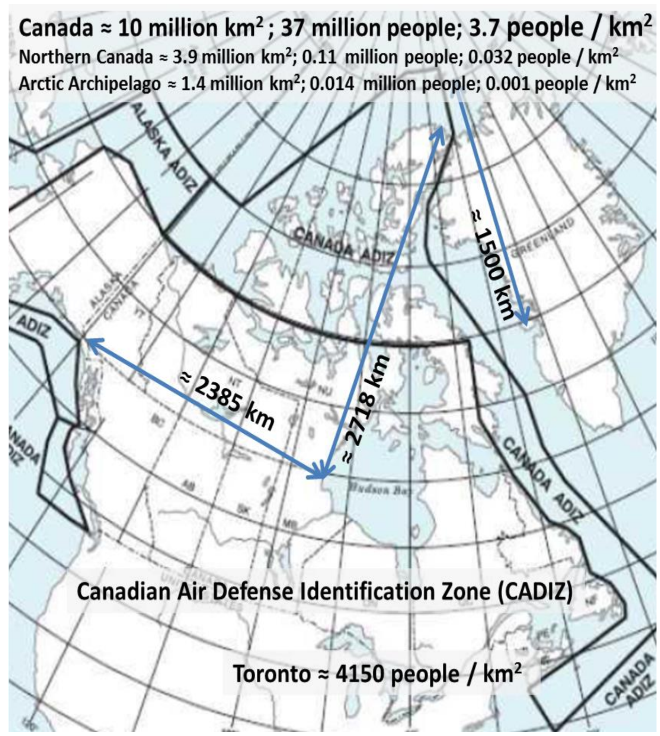

Figure 1. North American Canadian operational areas, new Canadian Air Defense Identification Zone (CADIZ), distance vectors and population densities for Canada, Northern Canada, Canadian Arctic Archipelago and Toronto.

Section II of this paper describes the geographic context, climate, topography and size of the area requiring assured communication in all space weather and atmospheric conditions. Section III summarizes relevant aspects of the communication systems envisaged, deployed and used over the last decades in this area, such as Tropospheric Scatter, Geostationary Earth Orbit (GEO) satellite relays at approximately 37000 km above the equator and Medium Earth Orbit (MEO) satellites between 2000 and 36000 km altitudes. Section IV introduces an option using a specific Highly Elliptical Orbit (HEO), the Three APogee (TAP with an apogee around 43000 km altitude over the CADIZ) [3] while Section V brings into perspective novel options offered by several Low Earth Orbit (LEO) satellite constellations at altitudes below 2000 km. Section VI proposes a possible terrestrial architecture with the addition of base stations using towers and/or Unmanned Aerial Systems (UAS) as gateways. Technical considerations begin with Section VII, which compares latencies among the different satellite systems' orbits and other communication technologies. Section VIII examines the pros and cons of some of these options. Section IX explores challenges experienced by radio channels in support of the variety of relevant operations with results from simulations to conclude the scientific analysis. Section X provides a short discussion in layman terms of the presented results and relevant suggestions for further considerations. Section XI concludes with a short summary of findings and recommendations.

II. Context

Northern Canada, aka the North, is the vast northernmost region of Canada variously defined by geography and politics. Politically, the term refers to three Canadian territories: Yukon, Northwest Territories, and Nunavut. This area includes the Arctic Archipelago and covers about 39% of Canada's total land area, and is home to less than 1% of Canada's population.

The Canadian Arctic Archipelago, more succinctly the Arctic Archipelago, groups together all islands lying to the north of the Canadian continental mainland excluding Greenland (an autonomous territory of Denmark). Situated in the northern extremity of North America and covering about 1420000 km^2 (550000 sq mi), this group of 36563 (named) islands in the Arctic Sea comprises much of Northern Canada—most of Nunavut and part of the Northwest Territories. This is about 15% of Canada's geographic area and is home to only about 0.04% of Canada's population. The Arctic Archipelago is experiencing effects of global warming, with some computer models estimating ice melting to contribute a 3.5 cm rise in sea levels by 2100.

CAF's preparedness training and exercises involve several thousand participants and observers, which temporally increase the population density in Northern Canada substantially. For example, Operation Nanook is an annual series of military CAF exercises in the Arctic. It is intended to train different CAF elements (Canadian Army, Royal Canadian Air Force and the Royal Canadian Navy) along other government organizations, such as the Canadian Coast Guard and Royal Canadian Mounted Police in disaster response training and Canadian sovereignty patrols throughout Northern Canada. Another series of exercises, Maple Flag, is conducted south of Northern Canada near CAF Base Cold Lake, which brings about 5000 participants. Both exercises place significant demands on information exchange interoperability, including: voice, data and video, with some telecommunications requiring low latency in order to fulfill machine-to-machine requirements.

A poignant factor in deploying communication systems in Northern Canada is the low population density, which does not support typical commercial business decisions to expand current telecommunications infrastructure the metropolitan Canadian population is accustomed to. With reference to Figure 1, population densities of Europe and India are respectively about 2.3 and 11.6 thousand times greater than those of Northern Canada. It is worth noting that some aspects of evolving 5G technologies are tuned to improve network capacity in high-user density areas like cities. Providing improved services in low density areas will require careful adaptation of current 5G strategies, standards and technologies for deployment in the North.

Summarizing relevant environmental and logistic conditions from [4], the North has extreme weather, with temperatures ranging from -50 to +20 degrees Celsius and wind gusts of up to (hurricane strength) 150 km/h. There is very little precipitation in the Arctic, with an average total of 100 to 200 mm of rain or snow per year. The amount of daylight varies with time of year and with latitude. In Resolute Bay, for example, the sun does not rise above the horizon from early November to early February and does not set from early May to late July.

Permafrost (perpetually frozen ground) is present in most of the Arctic and although typically roughly 10 m thick, it can extend 1 km below the surface. Construction of stable platforms required for large satellite ground stations or microwave towers is costly because the necessary support pillars must be driven below the permafrost; otherwise the platform will shift as the permafrost partially melts during the short Arctic summers (for non-scientific audiences, permafrost shifts like glacier flows). This requires careful site selection to ensure placement where the permafrost is acceptably thin to enable installation preferably into the solid ground underneath. Furthermore, this type of environment is similarly challenging for underground fibre installation.

Year-round access to all communities high in the Arctic is by plane only, increasing the overall cost of travel and accommodation for arctic operations and when planning and installing communications equipment. Communities use seasonal sealifts to transport non-perishable dry cargo (e.g., construction material, household goods, vehicles, etc.) and bulk fuel to them once or twice per year. Large or costly items to transport via air cargo, for example: large aperture satellite dishes or microwave towers, and associated construction materials, must be shipped by sea. This extends the logistic phase of any communication equipment deployment and/or exercise.

III. Practical Arctic Communications Systems

This section summarizes relevant aspects of the following communication systems: Tropospheric Scatter, High Frequency Ground Wave (HFGW), Point-to-Point (P2P) backhaul radio links, Fiber Optic Cable (FOC) and GEO satellite relay systems. Skywave propagation modes are excluded due to their susceptibility to space radiation and atmospheric changes.

Tropospheric Scatter, or Troposcatter, is a beyond-the-horizon communications solution using microwave signal scatter propagating through the troposphere. This phenomenon allows signals to be successfully received around the curvature of the Earth without direct line-of-sight between the transmitter and receiver over vast distances, up to 500 km. Troposcatter systems have been in use for several decades and the technology has seen advances in recent years, including improvements in throughput to 20 Mbps. However, for applications in the North it was found to be too expensive to sustain due to its power demand and infrastructure.

HFGW was documented as an alternate communication network for "nuclear-survivable means of communication for land-mobile missile systems in Europe" [5]. HFGW appears to be a potential alternative means of communication in case of satellite communications disruption, for example due to solar activity or other manmade disturbances. According to International Telecommunication Union (ITU) [6] there are different uses of the terminology and the surface wave is often called the ground wave, or sometimes the Norton ground wave or Norton surface wave, after Norton who developed tractable methods for its mathematical treatment. The generic formulaic expression for the ground wave is the sum of a direct wave, a reflected wave, and a surface wave. When transmitting and receiving antennas are close to the ground, the ground reflection coefficient is -1, and the direct and reflected waves act to cancel each other out, leaving the surface wave as the dominant component. Under such conditions, the ground wave is essentially equal to the surface wave. Empirical results [5] using broadband discone antennas with a cut-off frequency of 19 MHz operating over 20 to 30 MHz near the Arctic Circle between Norway and Germany showed good link connectivity for voice and data communications using narrowband channels for paths over irregular terrain across distances ranging between 19 and 115 km. Based on the empirical results reported, a communication system with its signal spread over 10 MHz with code division multiplexing and sufficient coding gain would offer throughput and fading resilience for high-reliability medium-data-rate channels.

Main challenges in providing microwave backhaul throughout the North using towers include the overall lack of infrastructure, support and staging, the inaccessibility of locations where towers would be built and powering such sites. Microwave links have been extensively used in the North providing reliable connectivity between communities in the Yukon and Northwest Territories that do not have direct access to FOC backhaul. Microwave links provide high throughput with low latency, a significant advantage over GEO satellite systems (below). As microwave frequencies require line-of-sight propagation, towers and topographic features are exploited to extend the range of links beyond the limitations imposed by the curvature of Earth and terrain features along the ground path.

An example of microwave technologies used in remote Arctic locations is the High Arctic Data Communication System (HADCS) of Ellesmere Island, which links Canadian Forces Station Alert (CFS Alert) at latitude 82.5° North (beyond line of site of GEO satellites) to Eureka over a distance of roughly 500 km. The overall communications path includes sending data via a GEO satellite link between Eureka (latitude 79.6° North) and Ottawa 4147 km away. HADCS was retrofitted in 2003 to run entirely on solar power, despite prolonged darkness during winter months. Integrated solar irradiance (W/m^2) over a certain time period and location is called solar radiation, solar exposure, solar insolation, or insolation (J/m^2 or kWh/m^2). Figure 2 shows a HADCS station powered by eight 120 W photo-voltaic (solar) panels arranged in an octagon (eight vertical panels distributed at 45° angular intervals to cover all azimuths).

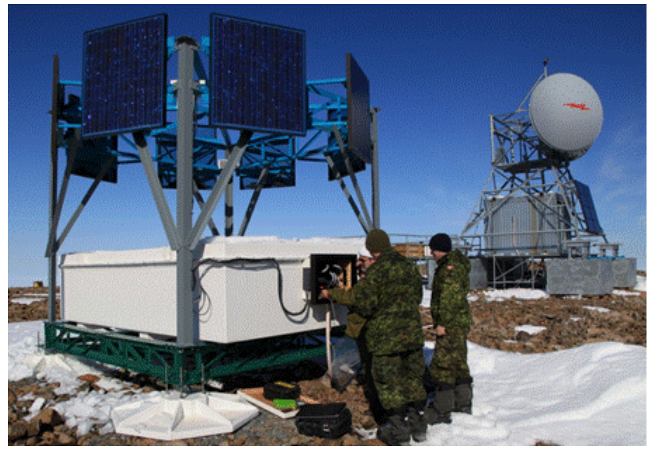

Figure 2. HADCS of Ellesmere Island.

During summer, with 24 hours of daylight, the sun does not dip below the horizon, and all of the solar panels contribute to charging battery banks [4]. Figure 3 shows CFS Alert's mean daily insolation values expressed in kWh/m^2/day being quite high (Miami, Florida has a value of ~5.26 kWh/m^2/day), making solar a viable option for CFS ALERT during the summer months with sufficient energy stored for the long winter.

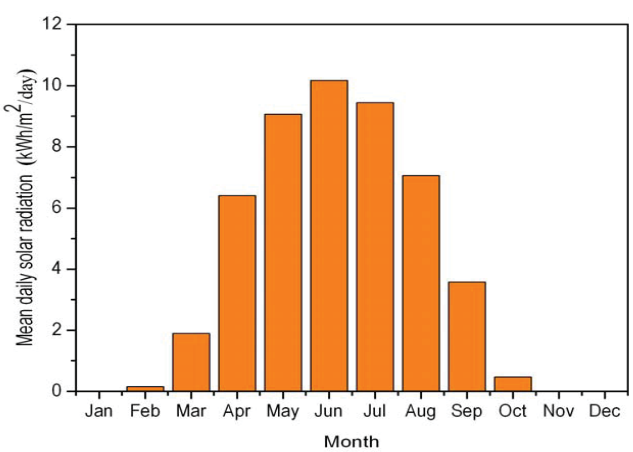

Figure 3. Variability of incident solar power for HADCS.

HADCS comprises seven individual links ranging from 18 to 121 km in length each. The current system operates in bands near 1800 and 2100 MHz, and provides 6.3 Mbps throughput. Helicopter transport is the only means of access to repeater sites [4] as shown in Figure 4.

In the North, costs of construction per kilometer for FOC and microwave links are about equal (approximately $65K/km), when serving a population of roughly 300 people. Above this number of end users, year round, fiber is less expensive [4]. These networks have at least one point of service provisioned via a satellite terrestrial terminal; in practice two or more are used to increase dependability.

Figure 4. A CH-146 Griffon helicopter slings a battery over a repeater site.

GEO satellite systems are well suited for some applications in the North. Given that GEO satellites remain at the same apparent point in the sky when viewed from a corresponding set of locations on Earth having clear line of sight, GEO satellite systems as a whole do not require tracking antennas on the ground as MEO and LEO systems do. This enables the use of lower-cost stationary antennas at ground stations, which is particularly advantageous in harsh environments where moving parts are undesirable. GEO satellite systems provide much of the broadband Internet coverage (including some backhaul connectivity) for communities in the North. Coverage in Nunavut is currently provided by two GEO satellites, Anik F2 and Anik F3, both using the C-band mainly between 4-8 GHz, but also using the 3.7-4 GHz range, which overlaps with the IEEE S-band [4].

However, the use of GEO satellite systems in the North has drawbacks. Due to geometry and the curvature of Earth, there is no clear line-of-sight between equatorial GEO satellites and locations above roughly 80 degrees latitude North as GEO satellites appear under the location's southern horizon. In practice, communities at latitudes higher than about 70 degrees have elevation-look angles to the GEO satellite below ten degrees as rays emanating from GEO satellite locations are essentially tangential to the ground. This leads to increased absorption and scattering of microwave frequency signals by the atmosphere more generically referred to as absorption, and as signal dissipation at higher frequencies due to precipitation (rain, snow and ice) more generically referred to as rain fade. The five northernmost communities in Nunavut are above 70 degrees latitude and the dominating factor is absorption as precipitation is low overall in the North.

Perhaps the greatest disadvantage of GEO satellite technology use for broadband applications is signal latency: due to the propagation distance to and from the satellite, the minimum delay for a round trip is 480 ms with a median latency of 600 ms. This is an order of magnitude higher than fiber and microwave link latency [4] (details below).

Tests of Mobile Satellite (MSAT) and Iridium capabilities for emergency communications in Northern Quebec showed insufficiencies. The Canadian MSAT system uses GEO satellites, which requires, at such latitude, a high-gain antenna and higher off the ground antenna mounting (resulting in larger installations, location constraints and thus greater costs). These conditions make (emergency) operations difficult and cost-ineffective with very little room to support emergency response.

The US commercial Iridium system with its current handheld, vehicle-mounted and fixed-remote equipment demonstrated different logistic problems, including high cost and poor performance of the handheld telephone sets, which incidentally usually cannot be used inside manmade structures [7].

The Iridium system was designed to be accessed by small handheld telephone sets, the size of a cell phone. Omnidirectional antennas, which are small enough to be mounted on such an Iridium phone, and the low battery power provided, are insufficient to allow the set's radio waves to reach a GEO satellite. In order for such a handheld phone to communicate with them, the Iridium satellites are at LEOs closer to Earth, at about 780 km above the surface. With an orbital period of about 100 minutes an Iridium satellite can only be in view of a handset for about 7 minutes, with the call being automatically "handed off" to another satellite when the previous one passes beyond the local horizon. This requires a large number of Iridium satellites in comparison with GEO satellite based solutions, carefully spaced out in polar orbits, to ensure that at least one satellite is continually in view from every high latitude point on Earth's surface. For seamless coverage at least 66 satellites are required, in 6 polar orbits containing 11 satellites each (with some inactive spares).

For a Canadian Arctic Underwater Sentinel Experimentation (CAUSE) project, transmission of data over a period of 7 months showed that the Iridium Pilot system (with an antenna about six meters above ground) did not fulfill required expectation [8]. "The Pilot data transfer rate for polar transceivers is far less than the advertised rate" [8].

IV. ENVISAGED HEO/TAP CONSTELLATIONS

Current GEO communication satellites leave the poles uncovered; consequently, Department of National Defence (DND) is exploring options of building the capability or acquiring services from commercial providers with future plans to cover this area. The initial operational capability is tentatively scheduled for 2029.

Quasi-geostationary coverage of Polar Regions can be achieved from HEO satellites. HEO/TAP [3] could be considered under the Enhanced Satellite Communications Project – Polar (ESCP-P) program to provide dedicated, secure and reliable Beyond Line-Of-Sight (BLOS) communications for domestic and continental CAF operations in the Arctic.

In accordance with Kepler's second law, a HEO satellite spends most of the time in the vicinity of apogee (i.e., the farthest point from the Earth's surface, farther than GEO). The orbit could be oriented in such a way that the apogee is over one of the two Polar Regions, so that only two HEO satellites can maintain a continuous view of (presence above) an entire polar zone [9]. When satellite A leaves the service coverage (optimal viewing) zone and heads toward the perigee (i.e., the closest point to the Earth's surface), satellite B rises over the same zone to maintain the same complete circumpolar region in sight. Interestingly, there are periods of several hours per day of coincident (i.e., stereo-like) viewing from the two satellites over most of the circumpolar service area. Such a system could provide meteorological imaging and communication capacity, similar to GEO (including high latency aspects), focused on the polar region. The first HEO satellite system with a period of rotation equal to 12 h called Molniya was implemented for communication purposes in 1965. It established that a two-satellite Molniya HEO constellation can achieve continuous coverage of the polar region 58°–90°N with a Viewing Zenith Angle (VZA) less than 70°. Another HEO system—with a 24-h period—called Tundra is currently used by the satellite Sirius XM Radio service operating in North America. Both orbits, 12-h Molniya and 24-h Tundra, require HEO satellites to be launched with an orbit inclination equal to 63.4°. This value called the critical inclination [9], corresponds to a zero rate of apogee drift due to the second zonal harmonic of the Earth's gravitational field, and ensures a stable position of apogee over the polar zone. If the HEO orbit inclination differs from the critical value, then the apogee gradually drifts toward the equator, requiring orbital maneuvers to maintain the intended orbit position. The farther the orbit inclination is from the critical value, the more resources are required to maintain the orbit. A drawback associated with the 12-h Molniya orbit is the risk linked to hazardous levels of ionizing radiation due to the satellite passing through the Van Allen belts. The highest danger originates from high-energy protons. The Molniya orbit crosses the proton radiation belts at the region of maximum concentration of energetic protons with energies up to several hundred MeV. As an alternative, a 16-h TAP HEO orbit was proposed, providing similar polar coverage as the Molniya HEO system while minimizing the proton ionizing hazards by extending the apogee to 43000 km [3]. The TAP orbit has a ground track with three apogee points repeatable over two days. Such a constellation of two satellites in TAP orbit still revisits ground tracks every 24 h.

V. PERSPECTIVE OF NEW LEO CONSTELLATIONS

Out of the 11 LEO satellite communications service proposals registered (2014 and 2016) with the US Federal Communications Commission (FCC), the following three are considered based on their maturity [10]: OneWeb, SpaceX (Starlink) and Telesat on Ku (12-18 GHz), Ka (27-40 GHz) and V (40-75 GHz) bands.

To ensure access to affordable high-speed Internet connectivity across rural and Northern Canada, the Government of Canada has invested $85 million and is committed to buy up to $600 million of services over 10 years following launch in 2022 of Telesat's LEO Satellite Constellation [11], which is leveraging Telesat's worldwide rights to ≈4 GHz of Ka-band spectrum.

The analysis in [10] is summarized as follows:

- The maximum total system throughput (sellable capacity) for OneWeb's, Telesat's and SpaceX's constellations are 1.56 Tbps, 2.66 Tbps and 23.7 Tbps respectively.
- A ground segment comprising 42 ground stations will suffice to handle all of Telesat's capacity, whereas OneWeb will require at least 71 ground stations, and SpaceX will need more than 123. And, in this respect, the considerations presented here above regarding ground stations are directly relevant for LEO coverage in the North, particularly satellite tracking antenna functionality.
- In terms of satellite efficiency (ratio between an achieved average data-rate per satellite and its maximum data-rate), Telesat's system performs significantly better (~59% vs. SpaceX's 25% and OneWeb's 22%). This is due to the use of dual active antennas on each satellite, and the lower minimum elevation angle required in their user links.
- OneWeb's system has a lower throughput than Telesat's, even though the number of satellites in the former is significantly larger. The main reasons for this are the lower data-rate per satellite that results from OneWeb's low-complexity satellite design, spectrum utilization strategy, orbital configuration, and payload design, and particularly the absence of Inter Satellite Links (ISLs) to which the analysis presented here will return later.
- If ISLs were to be used in OneWeb's constellation, (even with modest data-rates of 5 Gbps), the number of ground stations required could be reduced by more than half to 27 ground stations.

"To conclude, our analysis revealed different technical strategies among the three proposals. OneWeb's strategy focuses on being first-to-market, minimizing risk and employing a low-complexity space segment, thus delivering lower throughput. In contrast, Telesat's strategy revolves around highly-capable satellites and system flexibility (in diverse areas, such as: deployment, targeted capacity allocation, data routing, etc.), which results in increased design complexity. Finally, SpaceX's system is distinctive in its size; although individually each satellite is not significantly more complex than Telesat's, the massive number of satellites and ground stations increases the risks and complexities of the overall system considerably" [10].

However, although the massive number of satellites and ground stations increases the risks and complexities of the overall system considerably, our experience showed that with appropriate intelligent/adaptive designs, this offers a desirable high level of redundancy and ensures some self-healing capabilities required for dependable critical communication systems supporting vital time-constrained activities and operations, both civilian and military. Overall, employing such massive numbers of satellites might provide room for multifactor synergies beyond the scope of this paper briefly mentioned in the Considerations section X below.

It is worth mentioning that LEO satellite data collected during the 2014-2016 period for the cited 2019 article [10] may have changed considerably from the initial application through the FCC given the fast evolution of regulations and separately of technologies. According to current knowledge, SpaceX made several new applications to evolve the Starlink system design to a multi-constellation with three different orbit layer patterns, in addition to ISLs and a massive number of ground stations to increase connectivity to existing terrestrial infrastructure and creating new infrastructures where none previously existed. However, the current early version of the Starlink satellites do not have ISLs but instead they use advanced antenna arrays [12]. Given SpaceX linkages with National Aeronautics and Space Administration (NASA) and United States Department of Defense (DoD) (e.g., relative to DARPA's Blackjack program) [13][14], it seems that a possible overarching objective is to deploy capabilities for persistent surveillance, especially to detect and track hypersonic cruise missiles, which provides confidence in considering further synergies in the emergency management and preparedness sphere.

It is worth noticing that Starlink's beta testing is progressing as reported recently by Canada Satellite [15] More about Starlink progress could be found on the web as per the following references [16][17], as well as reducing interference with space observation with SpaceX's Darksat [17] which exhibited a 50% reduction although much more is required to fulfil the requirement for larger telescopes.

Several developments since the reference publication [10] affect the implementation space. Partly due to COVID-19 and loss of high-risk funding partners, OneWeb went bankrupt in March 2020 while trying to build a satellite constellation to deliver broadband. According BBC news, the UK is part of a consortium with India's Bharti Global, which won a bid to take the company over. Business Secretary Alok Sharma said it would help deliver the "first UK sovereign space capability". The situation is slightly different for Telesat's constellation, which intends to sell part of their licensed radio spectrum in order to generate funds for building their sophisticated LEO satellite technologies with commitments from the Canadian Government [11].

VI. TERRESTRIAL ARCHITECTURES

UAS gateways (either aerostats, hot air balloon, buoyant gas air balloon, tethered or free-flying, unpowered or powered, dirigibles or high-altitude high-endurance autonomous drones) [4] provide possible communication solutions that merit discussion. The Internet.org consortium has conducted some research into the feasibility of using UAS as communication platforms for remote and underserved locations [15]. Such UAS would be deployed at an altitude of approximately 20 km. By using solar power, UAS systems would be capable of maintaining station above a geographic location, thereby reducing complexity and cost of ground infrastructure when compared to microwave links, without requiring active tracking by the antenna on the ground typically used for MEO and LEO satellite systems. As the UAS would be relatively close to ground, cheaper low-power transmitters could be used, while still enabling high-throughput communications with low latency. UAS would be more susceptible to atmospheric weather than satellites. A previous study [4] assessed that this type of UAS range extenders might have an availability up to 80%. With intervening technology advances, such UAS might be sufficiently reliable today for commercial broadband.

Architecture options offered by new LEO satellite constellations and terrestrial communications, such as UAS, FOC and HFGW given advances in signal processing, multi-beam antennas, spatial diversity and low cost software-defined radios have the potential to substantially improve telecommunication systems availability and reliability in the North. Figure 5 illustrates a hybrid-technology architecture where ISLs and Inter UAS Links (IUASLs) play important roles. Long endurance UAS could use solar with hydrogen fuel-cells. UAS requiring refueling every six months would be ideal (north passage shipping season, tourist season, etc.).

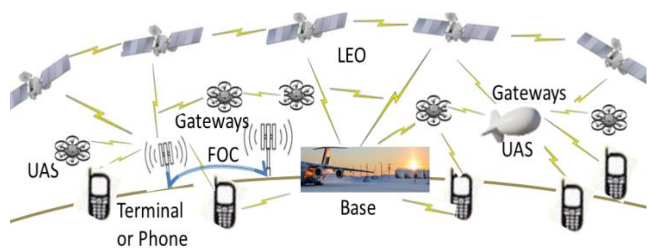

Figure 5. Simplified proposed Northern Communications Architecture.

VII. LATENCY

Now turning to specific technical aspects, the one way propagation delay of a message is the amount of time it takes to reach its destination from its source. Latency is the delay, usually measured in milliseconds (ms), that occurs in a round-trip data exchange. Round-Trip Times (RTTs) for FOC in large networks [18] using Content Delivery Network (CDN) show useful latency around 18 ms over a distance of 1400 km. HADCS and HFGW offer low latency in this order of magnitude at lower data rates.

High latency typical of GEO satellite links can be very disruptive for some applications, such as video conferencing, and could increase risk in remote health delivery applications including emergency response and particularly remote surgery. Satellite link latencies can also cause low data throughput, caused by the default behavior of communication protocols, which are optimized for shorter distances. A GEO satellite one-way propagation delay is approximately 240 ms due to the large distance between Earth surface location and the satellite; round-trip delivery of a data packet with acknowledgement is approximately 480 ms. This does not include a network delay, which can generally add 50 to 200 ms, depending on where the server is located. GEO satellite systems have a median latency of nearly 600 ms, which includes a median delay of 120 ms incurred by equipment processing speed and network delays in both directions. This makes GEO systems unsuitable to replace cable or fiber systems for applications requiring low latency, particularly impeding machine-to-machine interoperation.

For HEO TAP, with an apogee of 43000 km, the round-trip time is increased by 16% over GEO, i.e., 558 ms or a median latency of 778 ms. Assuming a MEO median orbit of 19000 km, that is 51% of GEO latency, then MEO median latency would be in the order of 306 ms.

The lower orbits of LEO satellites, however, result in latencies much closer to landline quality. The average orbit of the proposed constellations is around 1200 km, an average round trip of 2400 km, incurring a latency of about 31 ms. This is 93.5% improvement over a GEO round trip latency of 480 ms. If the processing speed of LEOs equals that of GEOs then their total median latency would be 151 ms. However, OneWeb tune-up, recorded an average latency of 32 ms in July 2019. As new LEO satellites are designed for high throughput, their overall processing time and network delay must be lower than those used in legacy GEO systems in order to obtain such low latency.

In this paper, since no large sets of empirical results are available for LEO latency, a conservative approach is to use selected simulation results from [19] for an optimal Expected Latency Minimization (ELM) algorithm used in the Software-Defined Networking (SDN) context, which addresses more comprehensively the overall network delay aspect including fading dependence on atmospherics. An interpretation of [19] for its ELM-SDN hypotheses is that LEO's average latency would be around 40 ms with a maximum average latency around 90 ms.

Considering the advantages of LEOs in extending telecommunications coverage and throughput of terrestrial network communications in the North with appropriate gateways, the analysis presented here is extended to include experimental findings.

One challenge of LEO constellations is the frequent handover (also known as handoff) between satellites or between their multiple spotbeams in a satellite footprint. Another challenge stems from terrestrial base station tracking antenna pointing limitations not limited to temperature dependent non-zero slew. These have a good chance to be mitigated for terrestrial base station equipment under low energy and cost regimes by the recent and expected availability of lower-cost components, such as high-speed low-power chip sets and Active Electronically Scanned Array (AESA) integrated boards for microwave systems in Ka and Ku bands. Use of phased array technology is common in satellite radio frequency antennas.

VIII. Pros and Cons of Some of These Options

Next we discuss some radio and LEO channel aspects.

A. *Common radio channel considerations*

Figure 6 illustrates generic terrestrial path loss as function of carrier frequency and shows disadvantages of transmitting at higher frequencies counterbalanced by higher frequencies offering higher throughput. Carrier frequency dependence is also present in satellite communications as mentioned herein.

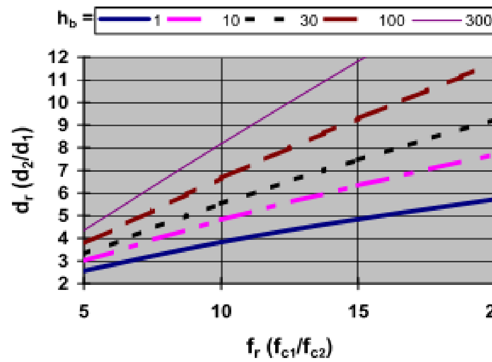

Figure 6. Path loss distance ratio d_2/d_1 as a function of the carrier frequency ratio f_{c1}/f_{c2} for five base station heights and mobile at one meter above ground.

Many of the problems of poor communications performance in remote areas can be alleviated by operating communication networks at lower frequencies. Radio-frequency (RF) passive and active devices, and ancillary components, are generally more efficient when operating frequencies are reduced from 10 GHz to 10 MHz. A similar tendency can be observed in propagation phenomena. Interestingly, urban noise increases as the frequency decreases but this is of less importance presently in the sparsely populated North. Overall, however, effective channel capacity and coverage at 150 MHz is superior to that at 1.5 GHz, at the same time a frequency allocation problem must be addressed (re-allocation of current VHF frequencies is required). The gain in energy-transfer efficiency at lower frequencies can be illustrated by using a propagation prediction model adapted by Hata [20], which estimates the power path loss L_p in dB as:

$$L_p = 69.55 + 26.16 \log_{10} f_c - 13.82 \log_{10} h_b - Ah_m + (44.9 - 6.55 \log_{10} h_b) \log_{10} d \qquad (1)$$

where the carrier frequency f_c is in MHz, antenna heights h are in m and the distance between the antennas *d* is in km. The effective base station height h_b range is from 30 to 300 m: however, in a deployable system for purposes relevant here, mobiles can be used as relays and the lower bound of h_b may be set to one meter. The correction factor $A(h_m)$ in dB is a function of the mobile effective antenna height h_m and size of the city (the Hata model dealt with an urban environment); we neglect this term in our model because the North is sparsely populated, in addition h_m is set to one meter. Applying (1) to two frequencies, f_{c1} and f_{c2} with a ratio $f_r = f_{c1}/f_{c2}$ larger than 1, all the other parameters of (1) being the same, and equating the path loss $L_{p1} = L_{p2}$, d_1 and d_2 become the dependent variables. With $d_r = d_2/d_1$, we obtain the following equation:

$$d_r = f_r^{\left(\frac{26.16}{44.9-6.55\log_{10} h_b}\right)} \tag{2}$$

Equation (2) indicates that, for a given path loss, we can increase the distance between the receiver and the transmitter by decreasing the operating frequency of the terrestrial wireless communication network, if other factors are unchanged. In rural areas, foliage and diffraction models for other shadowing effects and surface over-the-horizon radio-propagation allow the derivation of similar equations.

Figure 6 allows estimating the increase in radio coverage when stepping down from 1500 MHz to 150 MHz. Reducing the mobile operating frequency by a factor of 10 extends the communication range by a factor of about 5 for a base station whose effective antenna height is 30 m. If the cell size were 5 km for normal service, it might adaptively extend to 25 km for an emergency temporary service, reducing the logistic burden of covering an area affected by a disaster. Currently, the frequency of 700 MHz is allocated for emergency in Canada.

Colman *et al.* [4] present microwave link systems examples. One is operating at 1.8 GHz and the other at 11 GHz, with similar radiating power. The system at the lower frequency offers a free space maximum range of 333 km while the other, at six times the frequency, qualifies for a free space maximum range of 30 km, which is 11 times shorter. However, the maximum effective throughput rises from 65.4 Mbps to 232 Mbps, which is 3.5 times faster.

Other considerations include the challenge of powering terrestrial systems in the North briefly presented above in Section III. Sources like solar and wind mill power could be combined with sodium-ion batteries, operable at low temperatures.

These considerations apply with due change in particulars to generic mobile units, including airborne, and equally apply to UAS providing cell phone tower functionality.

B. LEO channel analysis

Equally, when a satellite is in direct line above a UAS, a ground station or a mobile transceiver, this is the shortest path, the Doppler effect is null, the signal is at its maximum level and usually is less affected by various atmospheric phenomena (cloud, rain, snow fall) aside from possible multiple ground or nearby structures' reflections. The signal is composed of a main direct ray and several secondary reflections. Such a channel displays fades statistics following a Rician distribution.

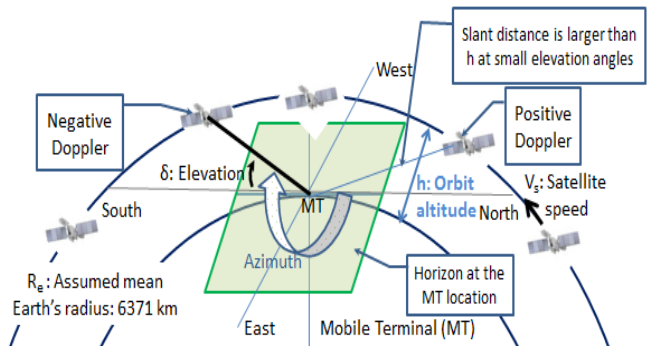

Figure 7. Moving ground or air terminal and LEO satellites.

However, when the satellite ascends (rises) above the horizon (positive Doppler shift) or descends (sets) toward the horizon (negative Doppler shift), i.e., at low elevation angles (δ) above the horizon as shown in Figure 7, the slant range path of the signal is close to the maximum useable distance before being completely shadowed by Earth's curvature. Note that polar orbit LEOs could rise from the North and later from the South over their complete orbit cycle since during a cycle the orbit path turn around Earth (East-West). In the following equations we assume that the mobile terminal is at a negligible height altitude. For UAS and aircraft at high altitudes these equations need to be modified. For a ground terminal, when δ is equal to zero, the slant distance sd can be expressed as follows:

$$sd = \sqrt{h^2 + 2hR_e} \tag{3}$$

with h = 1200 km and Re = 6378 km, sd = 4100 km. In order to ensure a sufficient signal strength level, LEO satellite constellation designers select a minimum δ large enough to ensure a useful workable Quality of Service (QoS). For example Telesat set δ to a minimum of 20°. Equation (4) [21] provides sd as function of δ:

$$sd = R_e\left(\sqrt{\left(\frac{h+R_e}{R_e}\right)^2 - \cos^2\delta} - \sin\delta\right), \tag{4}$$

which for δ = 20° gives an sd of 2453 km. Equation (4) can be verified with δ = 90°, which corresponds to the shortest distance value, i.e., sd = h as expected where the Doppler shift is null. Free space Line of Sight (LoS) path loss for each of these distances at 50 GHz (wavelength of 6 mm in the V band) are respectively 198.6 dB, 194.2 dB and 188 dB. Doubling range distance requires a four-fold increase in power for a 6 dB path loss increase at 20° of elevation. It is worst at 0° with about 10 dB or a power ratio of 10 just for the free space loss. However, as the elevation angle δ decreases, more adverse atmospheric phenomena will add to the total path loss, e.g., a cloud may add another 7 dB loss at that frequency.

In addition, the weaker (faint) received signal shows more Rayleigh fading, and is more prone to adjacent channel and jamming signal interference. At small elevation angles, sd is larger, the signal level is lower; the main path competes with relatively strong multipath reflections in addition to being more exposed to atmospheric phenomena over longer distances. Such a channel would likely display deeper and more frequent amplitude fades departing from the Rician distribution, however, following a Rayleigh distribution as hypothesized in [22].

Amplitude measurements of the received radio signal reveal time-varying characteristics resulting from propagation phenomena. When contributions to the total received energy arrive from a large number of reflections, giving a uniform distribution of phases each with similar amplitude, the resulting signal displays Rayleigh amplitude statistics. If a single contribution dominates, the total signal displays Rician amplitude statistics. When contributions arrive from only a limited solid angle around the receiver, the amplitude may follow either a Weibull or a log-normal distribution. In real situations, constant configuration changes as scenarios evolve lead to amplitude statistics that may vary considerably. For example, to address this variety of conditions, the Loo distribution offers an adaptation as function of the elevation angles [23].

The speed of a LEO satellite relative to a fixed ground station V_s can be expressed as function of its orbital period [24] as follows:

$$V_s = \sqrt{\frac{\mu_L}{h+R_e}}. \quad (5)$$

where μ_L is Kepler's constant 398600 km^3/s^2. Consequently the Doppler shift f_{ds} as function of the elevation angle and the carrier frequency [24] is:

$$f_{ds} = \sqrt{\frac{V_s R_e f_c \cos\delta}{C(h+R_e)}}. \quad (6)$$

Using (5) and (6) we find that $f_{ds} = 1059$ kHz for h = 1000 km, V_s = 7.35 km/s, and f_c = 50 GHz. These values represent significant challenges for data links relevant in this present study.

The available time to an initial connection to a LEO satellite, or one of its spotbeams, could be shorter than 3 minutes, hence the frequent need to access a newly visible channel. To address this, in view of challenges with Iridium presented in Section III above, there are a variety of handoff management approaches reported in [25][26]. Some address land mobile systems using terrestrial base stations. Here, we are more concerned with satellite-to-satellite handover [27] and between spotbeams, both inter and intra-satellite spotbeam handovers [28] serving terrestrial fixed or mobile users, base stations, gateways and UAS.

Some definitions from [29] for typical cellular deployments are reused for relevant purposes here. "A hard handover is one in which the channel in the source cell is released and only then the channel in the target cell is engaged. Thus the connection to the source is broken before or 'as' the connection to the target is made—for this reason such handovers are also known as break-before-make. Hard handovers are intended to be instantaneous in order to minimize the disruption to the call. When the mobile is between base stations, then the mobile can switch with any of the base stations, so the base stations bounce the link with the mobile back and forth. A soft handover is one, in which the channel in the source cell is retained and used for a while in parallel with the channel in the target cell. In this case the connection to the target is established before the connection to the source is broken, hence this handover is called make-before-break. The interval, during which the two connections are used in parallel, may be brief or substantial. Soft handovers may involve using connections to more than two cells: connections to three, four or more cells can be maintained by one phone at the same time. The latter is more advantageous, and when such combining is performed both in the downlink (forward link) and the uplink (reverse link) the handover is termed as softer. Softer handovers are possible when the cells involved in the handovers have a single cell site."

Problems with hard handover are the possibility of lost packets, lost requests to repeat them, cost of packet resequencing and consequently additional delays. Soft handover may monopolize more channel resources while being seamless to end users. As reported in [29], there are advanced soft handover approaches that better optimize retention of channel resources and reduce latency.

IX. RADIO CHANNELS TO SUPPORT OPPERATIONS

This section is based on the author's work [30], to which unpublished material is added illustrating the challenges of digital communications between collaborating entities for deployed operations. It illustrates how difficult it is to correctly assess the effective channel capacity in the context of several concurrent operations. The channel capacity is the most stringent factor affecting the performance of distributed information systems made up of mobile-computing nodes communicating via digital radios [31][32][33][34]. In the following analysis it is assumed that the distributed algorithms used are optimal for certain operational loads and observe that optimization is rapidly lost below a certain information exchange threshold [35]. To estimate the performance of such distributed algorithms accurately, several analytical and simulation models supported by experimental results have been proposed in [36][37]. In the majority of cases the models show that spatial and temporal statistical distributions of interrelated phenomena cannot be replaced by mean values without leading to large errors.

In operations such as SAR [38], emergency evacuation and forest-fire fighting, mobile units can be aircraft, helicopters, unmanned vehicles [39], trucks, all-terrain vehicles and backpacks. For this section we assume that information is shared via the radios in order to develop a common operational picture via a distributed database. Each mobile unit automatically reports its position to other participants on pre-established schedules. Decision makers coordinate operations with the aid of computers (accumulated information, geodisplay, decision support) and send appropriate control messages to participants via respective participant communication network nodes. Digital voice competes with the transmission of computer data.

For mobile computing, where purely digital data are exchanged over channels subject to Rayleigh fading, the effective channel capacity (throughput) decreases with an increase in relative velocity between communicating nodes for certain combinations of signal modulation, bit coding and error control protocols. This phenomenon is expected to be exacerbated when applied from terrestrial deployments, where relative velocities are at worst in the 100's km/h, to LEO satellites and UAS gateways where relative velocities are at least an order of magnitude greater.

Next, simulation results are presented illustrating the need to take into account the range of relative velocities encountered between communicating participating mobile computing nodes in order to estimate the performance of the associated distributed systems accurately. Without dismissing other challenges such as tracking the variation of the carrier frequency due to high maneuverability, speed and acceleration of the mobile computing node platform, especially when the Doppler shift changes sign, we focus on the problem of what to do when packet errors occur due to fast-fade phenomena, despite average Signal-to-Noise Ratios (SNRs) being adequate.

We investigate the performance of two different error-control schemes combining certain error-control techniques. The basic scheme uses error detection in a Selective-Repeat (SR) Automatic Repeat-reQuest (ARQ) scheme, SR-ARQ [36]. The other, a hybrid error-control scheme, adds forward-error correction (FEC), SR-ARQ/FEC [40]. For both schemes, Rayleigh-fading channels are assumed for participating nodes moving at relative velocities between 5 and 500 km/h (100 to 1 000 km/h in [41]).

The channel access scheme, a roll-call polling [42][43], is an energy efficient scheme employed in some distributed mobile applications including NATO tactical data exchange LINK-11 and is used here to illustrate aspects of the error-control protocols when sharing fast-aging information. It assumes that a master communication node is trying to gather information from other coordinated participating communication nodes and is the only node controlling channel access. Coordinated nodes wait until they are requested to send their updated data; if they have none, they return only a control packet with an acknowledgment (ACK). If no error-free response is received by the master node after a predefined delay, it polls the next participant node in its list. Reasons for no response include: 1- loss of radio connectivity to the intended participant node, 2- loss of radio connectivity from the participating node to the master, 3- collisions due to a previous loss of radio connectivity, and 4- radio silence either imposed at the queried participating node's platform or due to a fault.

A. *Motivation*

Assuming that many wireless systems will continue to exchange data using existing deployed radios with appropriate modified modes of operation, and/or internal or external upgrades, until new high-performance digital radios become more affordable, it is appropriate to present simulation and analysis results that point out the crucial tradeoffs to be made in link error control, to make best use of the scarce mobile-radio bandwidth. The effective channel capacity and error rate are highly variable and dependent on a variety of environmental factors. Such modified modes of operation, and/or internal or external upgrades, include interfaces to current radios, channel access schemes, error control, signal monitoring, automatic position reporting and interfaces to application computer systems. The selection presented takes into account recent interoperability trends within the telecommunications industry, where two types of circuit-mode services are used for computer-to-computer file transfer [44]:

1. Nontransparent: this service employs a radio link protocol that protects data during the mobile radio-transmission segment (as opposed to transmission over other media such as cable and/or satellite) including ARQ, FEC and flow control. Because of variable conditions in the mobile segment, effective user channel capacity decreases and delay increases as more packets received with errors must be retransmitted to maintain transmitted data integrity. Radio-channel bit-error rate is around 10^{-2} while most applications require data with bit-error rates better than 10^{-6}. This demands adherence to tight requirements on signal modulation, FEC, ARQ and flow-control combinations of such systems.

2. Transparent: this service employs FEC exclusively, i.e., without flow control or ARQ. Users must pre-establish a communication data rate and delay.

Research and development efforts for Asynchronous Transfer Mode (ATM) architecture suitable for mobile computing (civilian or military) may offer an alternative.

Figure 8 shows a model of the signal radiated by a mobile transmitter moving at velocity v_2 and affected by multiple reflectors, scatterers, diffraction layers and other propagation effects (e.g., free-space loss and shadowing), reaching another mobile receiver moving at velocity v_1. When a transmitter moves, its forward wave experiences an increase in carrier frequency and its backward wave a decrease; and at a stationary receiver corresponding positive and negative Doppler frequency shifts are experienced, respectively ($\Delta f_1 = \pm v_1$ / carrier wavelength). Similar effects occur due to a moving receiver: in the direction of the incoming wave front induces a positive Doppler frequency and vice versa. Because of multipath propagation due to scattering, diffraction and reflection, each received signal is represented by a complex vector sum (amplitudes, phases and Doppler shifts). In this paper we assume that the equivalent maximum Doppler frequency to be used in the Rayleigh fading model is the sum of the maximum contributions due to the two mobile platform transceiver velocities v_1 and v_2, that is $\Delta f = \Delta f_1 + \Delta f_2$. By extension, for LEO satellites communicating with a moving air or ground platform, the spreading of the received signal is limited to this value. For this paper we assume that the fade rate depends on the scalar sum of the two mobile (node) transceiver platforms' velocities ($v_1 + v_2$) relative to ground.

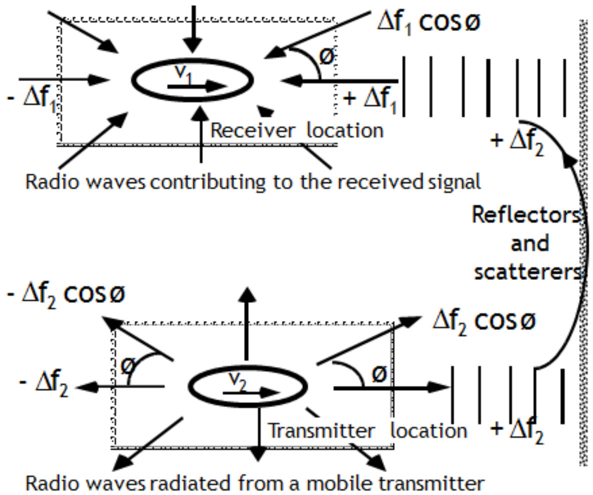

Figure 8. Transmission and reception of wave contributors for moving ground or space platforms.

Free-space path loss, clouds, (rain and/or snow) precipitation shadowing and slow fading, as well as the total noise, determine the local mean value of the observed instantaneous SNR. In the presented model, it is assumed that this SNR local mean is constant over the time required to transmit either: one bit, word, code word or packet. Path losses beyond the Gaussian reference are assumed to be due to fast fading following a Rayleigh distribution.

In [30], we selected a Rayleigh fading model for the fast-fade process, which is not limited to the case, in which adjacent bit errors are independent; the simulation also computes dependent bit errors when the in-fade period overlaps several bits. It is worth noting that the average fade duration is inversely proportional to the combined speeds of the mobile nodes, so it is also related to the Doppler shift. In addition, the number of bits affected depends on the bit rate. In most practical cases this distribution of fades causes higher word-, cell- or packet-error rates than if fades followed a normal, Gaussian or Rician distribution [45][46]. This model also assumes a mobile unit communicating to a base station, moving gateway, including UAS or LEO, or another mobile unit; consequently, the fade rate depends mainly on the mobile unit velocity, or both velocities, and the receiver local mean SNR. In such a case, given constant signal mean and threshold values, the time intervals between fades follow an exponential law and the time duration of fades obeys a Rayleigh law.

B. Error Detection

For most coding schemes we can use the following identification standard: (c, k) NAME-ID, for coding k information bits into c code bits with algorithm NAME-ID. In coding theory, notion of "perfect codes", which for binary repetition codes, are capable of correcting 0 to (c - 1)/2 errors, i.e., (c, k = 1) codes with only one bit of information since k = 1 (other "perfect codes" include the unique three-error-correcting (23, 12) binary and two-error-correcting (11, 6) ternary Golay codes). In general, these "perfect codes" are not suitable in terms of communication efficiency with unacceptably low code rates (1/c, for binary repetition codes). Error detection in our simulation is assumed to be perfect, i.e., all erroneous streams of c bits are detected. This is not the case for a real-world decoder, whose error-detection capability is limited: for large bit-error rates, the probability of undetected erroneous codes is not negligible. The assumption of perfect error detection suffices for practical purposes here, however, and is current practice in network simulation except when the performance of error-control techniques must itself be assessed specifically.

From simple information theory considerations, for independent error probabilities of adjacent bits, the code-word-error rate [45] without error correction is:

$$P(t=0, e=0) = 1- (1-p_{be})c \tag{7}$$

In (7) parameter t is the number of errors that a decoder can correct and e is the maximum number of errors acceptable in a received code word of c bits to deliver a correct code word: for decoders without error correction both parameters are zero. The channel bit-error rate is p_{be}. For Binary Phase Shift Keying (BPSK) over Rayleigh fading with 30 dB SNR, p_{be} can be around 0.0003 and for a code of 330 bits, P is 0.006 [46].

In our simulation, perfect error detection capability means that the decoder delivers a correct diagnostic with probability of one for any error pattern even when e = c.

The encoding scheme maps k user-information bits onto code-words of c bits, with c > k. The c-bit streams are designed to be more distinguishable among themselves in noisy conditions than the original k-bit streams, but this encoding increases the necessary transmitted bit rate by c/k.

C. Error Correction

When error probabilities of adjacent bits are independent, the code-word-error rate [45] with an error-correction capability of t code-word bits is:

$$P(t, e \le t) = 1-(1-p_{be})^{c} - \sum_{e=1}^{t}\left[{c! \over e!\ (c-e)!}\right](1-p_{be})^{c-e}\,p_{be}{}^{e} \tag{8}$$

We maintain the e = c condition for perfect-error detection. However, the condition on e for the decoder to deliver correct-code words is e ≤ t. Because our simulation considers the possibility of non-independent error probabilities of adjacent bits or of any set of bits in a code word, we cannot use (8). However, the conditions e ≤ c for detection and e ≤ t for error correction are valid.

Code-word errors are generated in the simulation according to the Rayleigh statistical fading model using a constant threshold [36]. Given perfect-error detection as indicated above, code words with t and fewer bit errors are corrected. Otherwise for e > t, the word is not corrected by the simulated error-correction decoder; and further actions to correct the word rely on the error-control protocols discussed in the next section. When e > t, real decoders induce new errors with a high probability; and usually the number of errors e' in the delivered word is limited as follows:

$(e - t) \geq e' \geq (e + t)$. We infer from previous work [36] that a word is in error if the sum e of all its bits affected by fading is larger than the decoder error-correction capability: $e > t$.

The highest code rate or code efficiency is 1; this occurs only without coding and overhead. Otherwise, code rates r are smaller than 1 with a number ($b_o > 0$) of overhead bits added to the number k of information bits to give a code-word length $c = k + b_o$. Consequently the code rate is $r = k / c$, and therefore $r < 1$. The bit rate after coding is $r_c = r_s / r = cr_s / k$, where r_s (in bit/s) is the rate, at which k uncoded information bits can be sent. Consequently, the bit rate after coding is larger than before coding for a constant transmission time, $r_c \geq r_s$. This means that if zero error is experienced by a message of k bits, then the message is received correctly at the sink without a need for coding; though that fact is not known in advance. Nevertheless adding coding means that $k + b_o$ bits must be transmitted instead of k; a loss in channel "capacity" proportional to the inverse of the ratio r, i.e., the inverse of the code rate without errors. For clarity, the bit-error rate indicates that only an average of $1 - p_{be}$ of the bits are correct, without the receiver knowing which ones are wrong.

D. Packet

A packet in our model can be defined as one code word (number of code words $n_c = 1$) or multiple code words ($n_c > 1$). The packet length is defined as $L_p = L_d + L_o$, where L_d is the total number of data bits ($L_d = n_c\, c$) and L_o is the total number of overhead bits ($L_o = n_c\, b_o$). More efficient coding can be obtained when error detection and correction are suitably selected, e.g., Viterbi decoding of short-constraint-length convolutional codes with a coding gain of 4.7 dB combined with a high code rate BCH code [46], which is beyond the scope of this paper.

The packet overhead L_o has three overhead components: L_{op} for the protocol, L_{od} for error detection and control, and L_{oc} for error correction alone. In this study, we maintain a sufficiently large L_o to satisfy the assumption of perfect error detection for the two types of packets used: data packets that contain user payload data, and control packets that contain information for node control, channel sensing and packet sequencing. The addition of training sequences may not be necessary since data packets are preceded by one control packet for the roll-call protocol.

According to the above definitions and choice of simulation parameters, we find that the code rate cannot exceed $L_d / (L_d + L_o)$. The code rate is further limited since L_o contains information about the data and some aspect of the previous state of communicating nodes. One advantage of having L_{oc}, L_{od} and L_{op} defined independently in the simulator is the ability to account accurately for overhead even when one of the parameters is set to zero, e.g., when t is zero, there is still some overhead associated with the error detection required by the selective-repeat error-control protocol. This accurate accounting of overhead is essential in measuring the effective channel capacity or user data rate: the amount of user information bits accurately transferred over a period of time in a given set of conditions, e.g., SNR in bandwidth, error-control techniques, modulation, channel-access protocol, environmental and operational parameters, and node load.

E. Protocols used

1) Error-Control Protocols

In the preceding section we discussed two means for controlling errors and information accuracy (code-word-error detection and correction) that are "forward-error-control" techniques, using the forward channel from a source to a sink of information. We further consider the possibility of a feedback channel from the receiver to the transmitter node, permuting these roles in order to feed back cues about transmission success(es) to the source. Error-control protocols are designed to use results of measurements made on immediate and previous observations of information accuracy and signal statistics at destination nodes. These results are then sent back to the source (feedback) where specific outcomes from the perceived accuracy and timeliness of delivered information allow automated counterbalancing actions to be triggered to correct observed situations; actions based on the feedback and based on a specific strategy. Here, error detection involves assessing the accuracy of the received message based on algebraic and probabilistic considerations.

Communicating units—the transmitter and the receiver(s)—store packets and node state variables using unique identifiers in order to distribute common references and to cooperate in the task of transferring user data from node to node correctly. When a packet is rejected by the decoder, an unambiguous request can be made for its retransmission.

For practical reasons (maximum allowable delay, buffer size and sequencing indices) the total number of outstanding packets must be limited. Raising this maximum increases the overhead needed to maintain coherent, unambiguous packet identification among participant nodes in the network. Also, it implies a larger total time to complete a full transmission cycle (the maximum delay in the case of correct delivery of a piece of information after some errors).

For tactical/fast aging data, the end users are generally more concerned with the most recent sensor measurement updates. Since both packet overhead and delivery time must be minimized for concurrent communication efficiency—especially for tactical/fast-aging data [35] over digital radios (less stringent for passive video playback)—as few as 11 bits can be used for sequencing. Other packet overhead bits are required for error detection and correction. Control packets contain a time-history of a receiver's perception of status/information and signal quality and in our study also convey channel-access information.

The resulting maximum value of latency depends on the maximum bit rate and maximum total number of outstanding packets. Practical considerations relevant here to fast aging data include a similarity between some types of tactical data and remote medicine delivery data in the sense that a most recent velocity (computed differences from measured locations) and a most recent heart beat rate (a time average) are important. In each application, the most recent measurement (GPS location / heart beat) can be absent, due to packet loss in transport, if the velocity / heart beat rate are statistically reliable. Receiver's perception can be slightly variable depending on whether such statistics are transmitted from source or computed at receiver. Conversely, in SAR operations all possible available real data is invaluable regardless of radio telecommunications network delay.

2) Selective-Repeat without Forward-Error Correction

ARQ schemes and protocols fall in the feedback-error-control technique family and are by far the oldest and most widely applied error-control protocols in use today [46]. ARQ schemes have three important sub-classes: Stop-and-Wait ARQ (SW-ARQ), continuous *per se* or Go-Back-N ARQ (GBN-ARQ), and continuous ARQ with selected repeats or Selective-Repeat ARQ (SR-ARQ). The last scheme, SR-ARQ, is more efficient with respect to effective channel capacity than the other two [45]. However, it is also the most complex of the three sub-classes.

We selected a basic SR-ARQ strategy with the previously defined perfect-error detector for one set of simulation parameters. The protocol always tries to identify problems with received packets; and a report identifying damaged packets is sent back to the sender. As is the case for any other transmission, this report itself can be disturbed during transmission over the radio channel, the sender may not be informed that a transmitted packet never reached its destination or that the ACK for this packet was disturbed on the return path. We did not limit the simulation to the noiseless feedback channel usually assumed in ARQ performance analyses [46] in order to account for more practical use situations.

In our protocol, we assume that a source tries to send a packet in its queue until the source receives, without error, a confirmation that the packet has been received correctly. At each transmission opportunity, a source node transmits N packets. If N_s of them are not perceived by the source node to be correctly received at the destination node, only $N - N_s$ new packets are taken off the source queue. This decision is made at every transmission opportunity, based on either no confirmation whatsoever, so $N_s = N$, or the correct reception of a control packet indicating N_s packets were determined to be in error by the destination node at its previous reception opportunity.

3) Selective-Repeat with Forward-Error Correction

Combinations of FEC with basic SR-ARQ strategies are called hybrid SR-ARQs. Some improved combinations use a decoding scheme employing multiple copies of a packet that were retransmitted because the first copy was in error. We elected to use FEC on a single packet copy, though this strategy is less efficient than the improved SR-ARQ/FEC, which employs multiple copies of a retransmitted packet (erasures and multiple copies) [37]; in most practical cases the selected scheme is less complex for only a small loss in performance. The important distinction is that in addition to perfect-error detection we add perfect-error correction, for up to t bits of a packet.

As previously, the protocol always tries to identify problems with received packets. If a packet has t or fewer errors, its error status is "no error in packet" and an ACK report is sent towards the source node. This ACK report is piggybacked in a Control Packet (CP in Figure 9) with some FEC, not necessarily a packet of the same length as a data payload packet. In general, control packets can be shorter than data packets. Although shorter packets are less likely to be damaged by an in-fade signal, they remain vulnerable during their transmission over the radio channel. As previously, the sender may not find out that one of its transmitted packets never reached its destination or that the ACK report was disturbed in transport via the feedback channel; with FEC on both types of packets, this happens less frequently than without error correction.

4) Representation of the Simulated Protocols

The protocols simulated are illustrated in Figure 9. We present their process flows by walking through the polling cycle over all the participant nodes (or participating unit labelled PU in the illustration), taking the following steps for each:

1. Master: *select* as current PU the next PU in the network.
2. Master: *read* the ACK from the previous transmission opportunity of this PU and *send* a request based on this ACK with an indication of the number of packets the master will send after this control packet.
3. Master: *send* its packet(s), if any.
4. PU: *send* a control packet including an ACK for the last master packet if it was received correctly, *do nothing* otherwise.
5. PU: *send* packet(s), if there are any and if a control packet was sent in the previous step.
6. Master: *listen* for a reply just after own transmission; if the PU's control packet is received correctly then *record* an ACK for this PU data packet (it may be a zero data packet); otherwise, after waiting one polling period, *record* a **null** ACK for this PU.
7. Master: *continue* cycling (optional end of cycle: *pause* to allow new PUs to join the net), and return to step 1.

F. Selected parameter values for the cases studied

We observed the packet-error rate (our dependent variable) as a function of the combined velocity of the communicating platforms (mobile computing nodes, UAS, LEO satellite), our main independent variable. Although, the simulation can estimate the effect of packet collision for the protocol of Figure 9, we did not select this mode for the results presented for concordance with the material presented in [47]. For similar reasons, we present participating node packet-error rates for a saturated network, that is: when the probability of having N packets in each transmitting queue converges to one or when the normalized offered traffic is equal to or greater than one. For conciseness, we do not present results for user message delay and effective channel capacity as functions of the relative velocity, since they are consistent with the results presented in [47]: they follow the packet-error rate.

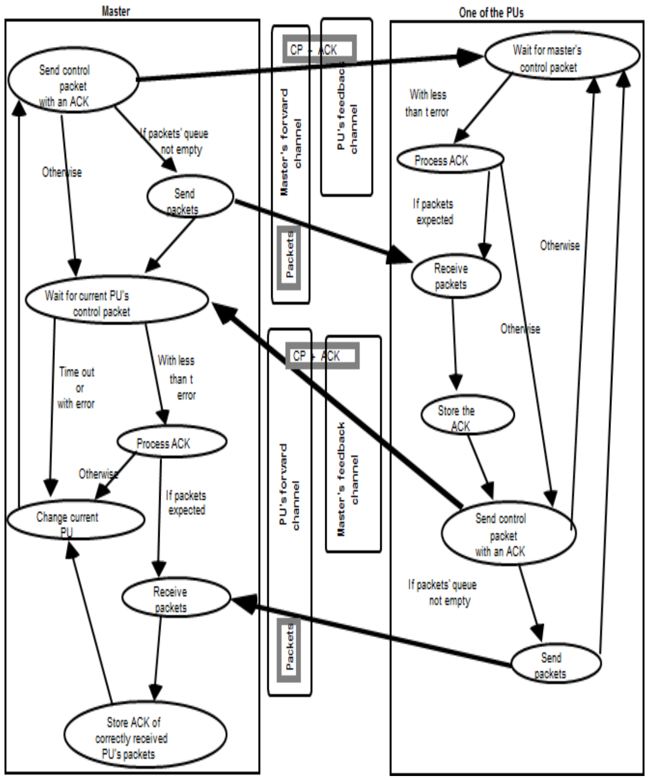

Figure 9. Selective repeat ARQ (with FEC, $t \geq 1$) with the roll-call control protocol.

We selected three operational control pairs of values of packet overhead L_o. The first pair simulates a network without error-correction capability, with L_o set to 30 and 20 bits out of 300 information and 200 control bits respectively. In Figure 10, results for this case are labeled t = 0. The two other operational control pairs correspond to networks with FEC capabilities. The simulation uses extraneous bits in addition to the 30 and 20 bits specified previously to increase the overhead by the required number of supplementary bits for correcting t errors: $L_{oc} = 2t$. In Figure 10, results for these cases are labeled t = 2 and t = 3 respectively. The two pairs of L_{oc} for error correction capabilities include one pair with 4 bits for data and 3 bits for control packets, which provide 2 bits and one bit of correction respectively, and another pair of L_{oc} for 6 bits and 4 bits, which provide 3 bits and 2 bits of correction for data and control packets respectively.

Using the selected Rayleigh fading model, high SNRs are associated with small values of the threshold-to-signal ratio ρ. Results for ρ = 0.001 are associated with a SNR of 30 dB in [47].

We assume binary modulation. The symbol period is $T_s = 1 / rc = 50$ µs for a transmission of coded bits at 20 kb/s. We set the values for $v_1 + v_2$ from 36 km/h to 936 km/h. These values are typical of: mobile computing platforms in range of a cellular communications tower or in radio communication range of a UAS. For a wavelength of 0.5 m the carrier frequency is 600 MHz and the maximum Doppler frequency ranges between 20 and 520 Hz. Since the mean value of the in-fade period can be longer than the symbol transmission period, bursts of errors are expected at low Doppler frequencies. The largest packet period is 16.8 ms, which most of the time is shorter than the no-fade interval for ρ = 0.001. For smaller SNR, packets will be affected by multiple fades (multiple bursts of errors); using larger symbols than binary will reduce this effect. The number of information packets N per error-control protocol cycle is set to three for the results presented. However, the model has been investigated by setting N as large as 20. The number of nodes including the master is three for the results presented. The roll-call cycle time has no influence on the results presented and will be used in evaluating the performance of the distributed system.

G. Information Packet-Error Rate versus the Effective Channel Capacity

We define the information packet-error rate (P_E) as the ratio of the number of retransmitted information packets relative to the total number of information packets sent including retransmissions. The normalized effective channel capacity or user's data rate (U) can be defined as the ratio of the number of correctly received information bits divided by the total number of bits transmitted:

$$U = \frac{(1 - P_E)(\mathrm{NL_{di}})}{\mathrm{N}\left(\mathrm{L_{di}} + \mathrm{L_{opi}} + \mathrm{L_{odi}} + \mathrm{L_{oci}}\right) + \mathrm{L_{dc}} + \mathrm{L_{opc}} + \mathrm{L_{odc}} + \mathrm{L_{occ}}} \quad (9)$$

The vertical axis on the right of the graph in Figure 10 shows the resulting effective channel capacity for t = 3, which corresponds to 1 - P_E multiplied by 0.73. With t = 0 the multiplier is 0.74. Consequently, there is no need to present separate graphs showing the effective channel capacities for the three multipliers since differences would not be noticeable. The dominant distinctions between the three error-control schemes are accounted for in P_E.

Given a fairly high constant mean SNR (30 dB), Figure 10 confirms the expected dependence of packet-error rates on velocity or Doppler frequency. This commonly observed Rayleigh fading channel condition for mobile radios is the most severe and difficult to overcome (except for Nakagami fading with a parameter smaller than one [48]): it usually causes the highest packet-error rate with a fade rate that increases with the relative velocities of the communicating nodes (mobile units, UAS, LEO satellites, gateways, towers, etc.) Similar results for fast-fading performance [49] support these findings.

In Figure 10, we also provide a typical result for packet-error rate assuming an additive white Gaussian noise (AWGN) channel model for the same SNR. This is a typical result from mean value analyses and valid for many combinations of noise, amplitude variation, modulation and coding. In this type of analysis the fading process is not velocity dependent. The AWGN performance measurements do not display the appropriate distribution statistics for mobile computing.

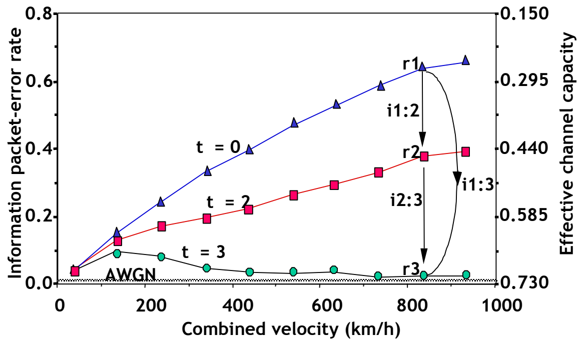

Figure 10. Selective repeat ARQ (with FEC, $t \geq 1$) with the roll-call control protocol.

Figure 10 needs to be adapted for higher carrier frequency, bit rate, packet length, code ratio and speed. If all these parameters are proportionally increased, one may observe similar results, i.e., when LEO is at low elevation, the mean SNR will be low (largest slant distance), with the perceived speed and Doppler near its maximum, this corresponds to higher error rates with no FEC (t=0) in Figure 10. While ascending to the shortest distance (equal to the orbit altitude) over a participating node location, a maximum SNR and lower Doppler shift will be attained, as for Figure 10 lowest speed.

H. Power Required for Equivalent FEC Performance

Through simulation we found that an increase in transmit power, labeled **i1:2** in Figure 10, required to bring the packet-error rate indicated by **r1** in Figure 10 to the level indicated by **r2** is **i1:2** = 5 (to decrease the error rate at **r1** to that observed at **r2** we need to increase the current power used by a factor of 5, e.g., from 5 W to 25 W). To obtain the same error rate without FEC, five times the power used with FEC that can correct two errors is required. The power increase, **i2:3**, required to decrease the error rate from **r2** to **r3** is **i2:3** = 332. The power increase, **i1:3**, from **r1** to **r3** is not equal to **i1:2** + **i2:3**; the requirement is **i1:3** = 1250 or from 5 W to 6 kW, a generally unacceptable value for typical long term operation of mobile and handheld units. However, it may be a desirable mode of emergency operation, for example in support of SAR, wherein a powered radio typically operating continuously is switched to a sporadic high power transmission mode to maximize use of an emergency power battery. These three values were estimated in the same way as all other point estimates used to produce the chart in Figure 10 and are consequently based on statistics on random variables. The relation between these variables seems to be **i1:2 • i2:3 = i1:3** with $5 \bullet 332 = 1660 \approx 1250$. With more samples of this nature we can develop an empirical model expressing the required power ratio as a function of system conditions and parameter values.

Increasing the average SNR by increasing the power of the transmitters to maintain a sufficient data rate is not practical for most mobile applications. A better approach would be to use additional coding and processing gain, e.g., FEC and power spreading. Here we assume that FEC acts to introduce time diversity and spread spectrum as frequency diversity, and both alleviate the impact of high fade rates on packet-error rates as relative velocities increase. For tactical operations, this can provide some level of resistance to jamming through coding to mitigate some level of signal jamming during adversarial attacks. However, with spread spectrum, one can spread the energy of the transmitted signal to a point where it is buried in noise, making it less detectable by basic/legacy electronic support measures systems, although still detectable by more advanced systems. The resulting data rate is higher at high relative velocities but slightly lower at low velocities, making the achievable effective channel capacity more nearly constant over a range of velocities.

The effective channel capacity penalty due to coding is merely a small decrease in the maximum data rate available at negligible relative velocities. Figure 10 shows the total energy saving achievable for the considered system if signal processing, 2 or 3 bit FEC capability, is used instead of higher transmitted power [48]. Using coding and spreading the baseband signal are techniques known to be implementation efficient. For lower SNR and to meet some mobile computing requirements, more coding and processing gain might be required than what is used in our examples.

I. Distribution of Packet-Error Rate Mean Values

We selected the point **r2** from Figure 10 to explore the smoothness of the stochastic processes simulated. As presented in the frequency histogram of Figure 11, there is indication of a good match to a normal distribution of the mean values (point estimates) of the packet-error rate. Note that this normal distribution of mean values of the packet-error rate computed from 600 independent simulation runs seems to be expressed naturally from the statistical processes at play. The temporal distribution of the packets experiencing error(s) cannot be deduced from these results. Accurate performance evaluation of distributed algorithms requires considering distributional dimensions of the underlying systems.

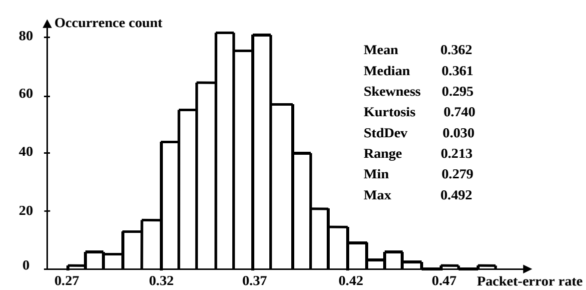

Figure 11. Statistical distribution of point estimates of the packet-error rate found from 600 independent simulation runs.

J. Asymmetry

The improvement for a small FEC capability increase on packets, from 2 to 3 bits, affects error rates experienced by all participant nodes including the master, as indicated by Figure 12. The improvement is larger at higher velocities because of a better match between FEC and error patterns. The in-fade period decreases and the rate of fades increases with velocity though the SNR is constant. The improvement is even larger for the master, since more control packets are involved.

This difference in the packet-error rates between packets sent by the master and by other participant nodes is quite noticeable. Since we selected to not account for collisions, higher master error rates are not due to the larger probability of its packets to collide with PU packets caused by the roll-call protocol asymmetry between a master and participant nodes, but are simply due to the conditional error probability based on the protocol asymmetry. Packets from a participant node are sent to the master only if the appropriate control packet from the master is received correctly, since we ignore bit streams after an incorrectly received master control packet that contains the master reception report concerning the last transmission made by a particular participant node. Consequently, the master reception report on participant node packets is sent on the master-to-participant feedback channel until it finally reaches the concerned participant node without error. The participant node does nothing but wait. For data packets sent by the master the scenario is different. The participant-to-master feedback channel for the reception report concerning master packets does not have this error-free property because each participant node transmits only when it has correctly received an invitation to send data.

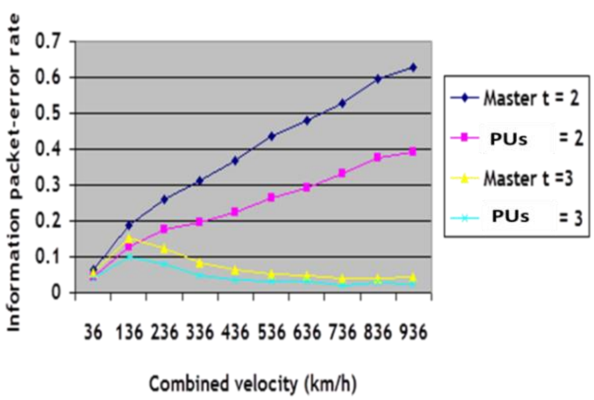

Figure 12. Master's and PUs' information packet-error rates for two values of t.

Despite relatively high average SNR, we observe that the effective rate of correctly transmitted packets drops dramatically as the velocity of communicating nodes increases. The fade rate increases with the velocity while the infade interval is shorter. As the SNR decreases the fade rate increases, but, with longer infade intervals. Both cause an increase in the probability of packet errors on the channel.

K. Summary of the simulation observations

Using mean-value analysis would have precluded noticing the effects of this decrease of effective channel capacity or increase of packet error rate with an increase in relative velocity between all participating nodes. The same is expected for performance analysis of distributed algorithms based on mean-value analysis; parametric empirical models will enable an analyst to visualize the effect of some tradeoffs.

The selected SNR, packet lengths and protocol overhead illustrate that if the design were based only on mean-value analysis, the resulting system would be insufficient to fulfill expected/user operational requirements. Although the effective channel capacity is generally adequate at negligible velocities, the effective channel capacity drops and becomes too small for practical use at relative velocities (ΔV) above 100 km/h. In air-to-air scenarios relative velocities may greatly exceed this range; supersonic speeds often exceed 330 m/s or 1200 km/h.

As the packet error rate increases, the power of the transmitter needs to be increased by a much greater factor to maintain the effective channel capacity and error rate. Typical area-coverage requirements, spectrum management, cost and technology constraints impose on designers the selection of alternatives that combine better modulation, error-control techniques, coding and baseband signal spreading.

In essence, the research results presented here above encapsulate a particular desirable operational parameter space for software defined radios.

X. Considerations

While the results presented here above address technical and scientific aspects of software defined radios, it is hoped that perhaps non-technical audiences find confidence in these results for situation aware development, deployment and coordinated interoperable equipment. In layman terms therefore, the solutions supported by the above analysis amounts to situation adaptable radios.

While costs for tactical operations/scenarios, preparedness exercises and border patrol can be justified in individual jurisdiction (Canada, NORAD), global solutions (NORAD, NATO) can be hard to justify. However, multinational corporations, such as SpaceX and OneWeb (new consortium), without excluding Telesat, are welcome intermediaries particularly in shared fate scenarios:

- Extending the reach of general peaceful operation of internet services to communities in the North remains dominated by commercial factors heavily influenced by population densities and the vast distances between/to Northern communities.
- While not fully sufficient for tactical/border patrol/emergency response/SAR and related preparedness exercises, GEO, MEO and HEO can and do support environmental monitoring and alerting, transactional banking, some provision for remote medicine virtual visits, triage and escalation, while perhaps insufficient for emergency response.
- While typical cellular communications enjoyed in Canadian metropolitan areas may not be available in the North, cellular mobile communications including mobile computing can be employed temporarily in the North in support of preparedness exercises, patrol operations and

prolonged emergencies by employing cellular tower equivalent equipment on airborne platforms including, but not limited to: patrol/SAR recon planes, aerostats (tethered or untethered), dirigibles and particularly on UAS.

- The results presented here particularly show sufficient support for mobile computing in the North via UAS gateways, which when used with the existing terrestrial infrastructure can incorporate, reuse and extend reach. Use of IUASLs can flexibly extend reach further and beyond geographic limitations, at least temporarily (tourist season, shipping season in the Northern Passage as the North Passage becomes more available both due to manmade advances and/or climate change without excluding broader uses) as well in justifying communication infrastructure build-out in the North.
- UAS can be further used based on the results presented here for interoperability as a middle relay layer between LEO satellites and terrestrial infrastructure via software defined radios, which account for high relative velocities between airborne UAS and LEO satellites in outer space. Essentially, by accounting for high Δv, UAS-to-UAS links become substantially equivalent to UAS-to-LEO links for the duration during which a particular LEO satellite is in direct line of sight of a UAS not unlike a cellular mobile set moving within a cell served by a cell tower. Both mobile computing node-to-UAS and UAS-to-LEO radios need to account for corresponding Δv Doppler shifts/fading in similar ways although it is understood that the radios on UAS serving terrestrial mobile communications below are configured differently than the radios enabling space communications above the UAS gateways with LEO satellites.
- Use of phased antenna arrays is promising not only in ground-to-UAS and particularly in direct ground-to-LEO communications to address environmental conditions in the North deleterious to antenna tracking to point to UAS/LEO satellites, but also in simplified UAS designs to handover UAS-to-LEO radio communication channels between UAS space pointing spot beams.
- While low altitude UAS gateway deployments can be susceptible to high winds, the treatise presented here shows balancing gains from reduced precipitation related signal degradation typically plaguing more southern radio communication infrastructure where precipitation is more abundant. Certainly, these results permit selection of UAS operating altitude and therefore retaining interoperability with legacy equipment.

Services provisioned with appropriate parameters and hardware anticipating improvements in coverage, resilience, redundancy, dependability, data rate and low latency include:

- fixed installations like CFS Alert, NWS radar networks, and Forward Operating Bases (FOBs);
- mobiles near fixed installations, airborne or tower gateways or their communication relays;
- short term deployed personnel and platforms for military exercises and operations, emergency operations; and
- off-grid communities of Northern Canada.

Also, there is a need to investigate how Canada could protect space and terrestrial network installation assets.

Satellite transmissions are more susceptible to radiation, jamming and atmospheric disturbances than FOC and over-the-horizon HFGW transmissions. HFGW at 20 to 30 MHz is expected to provide reliable medium throughput for terrestrial communications [5]. FOC offers high throughput and low latency, is commonly deployed around the world and is expanding in Northern Canada [50][51]. However, FOC and HFGW do not offer the area coverage of LEO satellites.

While national direct investment into LEO full constellation deployment remains hard to justify as each of large numbers of LEO satellites spend such a little time over any particular geographic area, the results presented here point to promising and accountable research and development investments.

Further research in mobile computing must be conducted to optimize the sharing of fast-aging data over radio networks (including satellites) used in planning operations critical to our communities, including search and rescue, ice-storm or flood evacuation, etc. Solutions to these problems must be global so that the various organizations involved in such operations—local, national and international—can respond to the needs of the threatened population promptly and cost effectively.

It is self-evident that the move from GEO satellite supported communications to LEO satellite supported communications is a paradigm shift in which:

- Extremely few purpose built, highly redundancy designed, long term operational life expectancy of GEO satellites using long term high cost development cycles are subject to catastrophic failures.
- Extremely large numbers of LEO satellites are built as generic as possible to serve multiple jurisdictions compliant with multiple jurisdictional standards, via software defined radios, for term operation following short term agile development cycles where redundancy is provided by sheer numbers of LEO satellites.

While redundancy has not been explicitly discussed in the technical development in this paper, it bears mentioning that just as the Iridium system makes use of passive pre-deployed Iridium satellite spares, LEO constellations rely on the use of LEO satellite spares. In SpaceX's case, SpaceX's quick low cost turnaround of reusable rocket stages demonstrated just recently, the LEO satellite spares do not have to all be deployed, but a few. The paradigm shift extends from previous hand manufacturing and assembly of GEO satellites to assembly line manufactured LEO satellites and therefore to 'relatively abundant' supply of LEO satellite spares.

Perhaps, beyond the focus of this paper, LEO satellite spares can be very interesting, both the large number of space deployed spares and on the ground LEO satellite spares awaiting launch. Certainly, it can be appreciated that due to low life time expectancy of space deployed LEO satellites, stand-by LEO satellite spares are wasted leaving one to ponder uses for active LEO satellite spares. While certainly, LEO satellite spares are required by the operator to protect investment for commercial ends alone, what is the value of an active LEO satellite spare orbiting over a jurisdiction experiencing a disaster? What is the value of such an active LEO satellite spare for public safety? What is the value of such an active LEO satellite spare in SAR situations regardless of geopolitical borders?

Equally relevant to immediately related fields, what provisions related to the activation of LEO satellite spares can be made in communication license approvals for LEO constellations? Would investment into LEO constellation operations be more palatable, as accountable, for example as insurance premium advance payments equivalent to a rocket launch for 10 to 60 replacement LEO satellites to relieve bandwidth commandeered during extended emergencies? In spite of OneWeb's recent financial troubles, this paradigm shift has made these questions possible.

While the treatise presented here cannot definitely solve all communications shortcomings in the North equally in all situations and in all scenario, UAS gateways are proposed as a middle layer for ground-to-air and air-to-air mobile communication nodes. The obvious question is what is the development and deployment cost and the time horizon of such UAS gateways? The developments presented here show an overlap interoperable not only with legacy equipment while bridging in LEO satellite radio communications in which UAS gateways are no different in radio transmission terms than a LEO satellite constellation layer except for being airborne under. Perhaps another aspect of the paradigm shift in the production of LEO satellites can be appreciated, the assembly line manufacturing can, and often does, produce satellites unfit for space launch due to manufacturing shortcomings of space operational essential components (punctured fuel tanks etc.) making such hardware available for integration into UAS where space essential components are completely unneeded. With this in mind perhaps investment into multinational LEO satellite development and constellation operations can be justified through national acquisition of select such LEO satellite production fractions, at least for research and co-development of UAS gateways.

Aside from tactical considerations supported in this paper, success of LEO satellite enabled communications can more comprehensively address remote medicine delivery in the North including emergency response, perhaps remote surgery. As this paper is being submitted, the implications presented herein are not just thinkable; SpaceX is gathering private interest in LEO internet testing direct to ground.

XI. Conclusion

This article addressed some difficult remaining challenges after many years of communications systems research and development for the North. Findings address DND/CAF challenges to improving capabilities required by future assured communications demands from expected developments in the Arctic and for NORAD operations.

Selected options to improve communications in off-grid areas, more specifically in the North, are expected to provide timely improved shared situational awareness in support of operations where it is currently not well provided, or not available. Such solutions would be revolutionary for our Defense and Security (D&S) capabilities and would progressively provide significant advantages to coalition forces and when CAF operates in collaboration with other Canadian departments including PS and local police in the most demanding emergency and disaster situations.

In the North, if low latency communications are critical for applications or operation objectives, GEO, MEO and HEO satellite systems are insufficient. The least expensive communications systems with low latency in the North should include microwave links, FOC, UAS and LEO systems to ensure fast deployment and access to a large majority of participants in most pressing situations. UAS offers rapid deployment capability on demand in response to PS and CAF situations. LEO/UAS hybrid systems could most definitely extend capabilities of available legacy infrastructure with microwave link and FOC terrestrial infrastructure ensuring connectivity with existing Northern networks and users.

As supported by the results presented here, for moving computing platforms (LEO, UAS, end user vehicle), system designers would have to consider appropriate protocols to match end user requirements for operations. Inevitably, a balance in processing gain, FEC and error control will need to be estimated to match the expected Doppler shift, minimum and maximum relative speed, fast change of in path length affecting free space path loss and atmospheric effects, slow and fast fading, and frequent handovers required.

Overall, the most significant finding is that the advent of low-cost high-performance LEO satellite systems, in conjunction with UAS, can substantially improve communications in the North supporting sustained research and development investment.

Acknowledgment

We are grateful to Marie-Josée Bureau (now at AEG in Montreal) who created the simulation programs critical to this research and provided stimulating and constructive discussions on the approaches and models. We also want to thank Nicholas Bond, one of my former research colleagues, (currently a consultant) for his editorial review, comments and contributions. Finally we thank DRDC for supporting this paper and my wife France Crochetière for her support during this challenging period imposed by COVID-19. Sensitive to current world events, some typical control theory language has been altered herein without loss of essence.

References

[1] P. Labbé, "LEO Satellite Constellations: An Opportunity to Improve Terrestrial Communications in the Canadian Arctic," in *Twelfth International Conference on Advances in Satellite and Space Communications (SPACOMM)*, Lisbon, Portugal, 2020: International Academy, Research, and Industry Association (IARIA), 2020.

[2] Canada: Department of National Defence and Canadian Armed Forces. (2017). *Strong, Secure, Engaged. Canada's Defense Policy*.

[3] A. P. Trishchenko, L. Garand and L. D. Trichtchenko, "Three-apogee 16-h highly elliptical orbit as optimal choice for continuous meteorological imaging of polar regions," *Journal of Atmospheric and Oceanic Technology,* vol. 28, no. 11, pp. 1407-1422, 2011.

[4] G. Colman *et al.*, "Communications in the North," Communications Research Centre (CRC), Ottawa, 2014.

[5] J. R. Champion, "An Empirical Investigation of High-Frequency Ground Wave Propagation," *Johns Hopkins APL Technical Digest,* vol. 13, no. 4, pp. 515-525, 1992.

[6] ITU, "Handbook on Ground Wave Propagation," in "International Telecommunication Union; Radiocommunication Bureau," R-HDB-59, 2014.

[7] P. Labbé, "GPS and GIS Integration in Mobile Equipment for Improved Mobile Emergency Operations," in *Proceedings of the 12th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-99); Session A2 - Land Systems & Public Safety*, Nashville Convention Center, Nashville, Tennessee, 1999, pp. 545-555, 1999.

[8] G. J. Heard, "Initial report of CAUSE investigated technologies for remote system deployment," Defence Research and Development Canada (DRDC) – Atlantic Research Centre, August 2019.

[9] A. P. Trishchenko, L. Garand, L. D. Trichtchenko and L. V. Nikitina, "Multiple-apogee highly elliptical orbits for continuous meteorological imaging of polar regions: challenging the classical 12-h Molniya orbit concept," *Bulletin of the American Meteorological Society,* vol. 97, no. 1, pp. 19-24, 2016.

[10] I. del Portillo, B. G. Cameron and E. F. Crawley, "A technical comparison of three low earth orbit satellite constellation systems to provide global broadband," *Acta Astronautica,* vol. 159, pp. 123-135, 2019.

[11] J. Foust. (2020, 11 November) Telesat remains optimistic about prospects for LEO constellation. *SpaceNews*. Available: https://spacenews.com/telesat-remains-optimistic-about-prospects-for-leo-constellation/, Access date: 12 November 2020.

[12] C. Henry. (2019, 15 May) Musk says Starlink "economically viable" with around 1,000 satellites. Available: https://spacenews.com/musk-says-starlink-economically-viable-with-around-1000-satellites/, Access date: 12 November 2020.

[13] S. Erwin. (2020, 11 July) Loft Orbital satellite to carry experiment for DARPA's Blackjack program. *Spacenews* [News].

[14] S. Erwin. (2019, 26 February) Air Force laying groundwork for future military use of commercial megaconstellations. *Spacenews* [News]. Available: https://spacenews.com/air-force-laying-groundwork-for-future-military-use-of-commercial-megaconstellations/, Access date: 14 Nov 2020.

[15] Canada Satellite, "Starlink Beta Testing Services Terms and Conditions," ed. Calgary, Alberta, Canada: Canada Satellite, 2020.

[16] M. Sheetz, "SpaceX prices Starlink satellite internet service at $99 per month, according to e-mail," ed. USA: Consumer News and Business Channel (CNBC), 2020.

[17] J. Tregloan-Reed *et al.*, "First observations and magnitude measurement of SpaceX's Darksat," *arXiv preprint arXiv:2003.07251*, 2020.

[18] I. N. Bozkurt *et al.*, "Dissecting Latency in the Internet's Fiber Infrastructure," *arXiv preprint arXiv:1811.10737*, 2018.

[19] W. Cho and J. P. Choi, "Cross layer optimization of wireless control links in the software-defined LEO satellite network," *IEEE Access,* vol. 7, pp. 113534-113547, 2019.

[20] M. Hata, "Empirical Formula for Propagation Loss in Land-mobile Radio Services," *IEEE Transactions on Vehicular Technologies,* vol. VT-29, no. 3, pp. 317-325, 1980.

[21] S. Cakaj, B. Kamo, V. Koliçi and O. Shurdi, "The Range and Horizon Plane Simulation for Ground Stations of Low Earth Orbiting (LEO) Satellites," *IJCNS,* vol. 4, no. 9, pp. 585-589, 2011.

[22] N. Okati, T. Riihonen, D. Korpi, I. Angervuori and R. Wichman, "Downlink Coverage and Rate Analysis of Low Earth Orbit Satellite Constellations Using Stochastic Geometry," *IEEE transactions on communications* no. tbd, 2020.

[23] S. K. Sharma, S. Chatzinotas and P.-D. Arapoglou, *Satellite communications in the 5G era*. Institution of Engineering and Technology, 2018.

[24] S. Mosunmola B, A. O. Agboola, A. Felix and A. Mohammed, "The Mathematical Model Of Doppler Frequency Shift In Leo At Ku, K And Ka Frequency Bands," *Published in International Journal of Trend in Research and Development (IJTRD),* vol. 4, no. 5, October 2017.

[25] I. F. Akyildiz, H. Uzunalioğlu and M. D. Bender, "Handover management in low earth orbit (LEO) satellite networks," *Mobile Networks and Applications,* vol. 4, no. 4, pp. 301-310, 1999.

[26] P. Carter and M. Beach, "Handover aspects for a Low Earth Orbit (LEO) CDMA Land Mobile Satellite (LMS) system," 1993.

[27] L. Wood, "Internetworking with satellite constellations," PhD, University of Surrey, 2001.

[28] R. Musumpuka, T. M. Walingo and J. M. Smith, "Performance analysis of correlated handover service in LEO mobile satellite systems," *IEEE Communications Letters,* vol. 20, no. 11, pp. 2213-2216, 2016.

[29] S. Das *et al.*, "Location Manager based Handover Method for LEO Satellite Networks," *International Journal of Computer Applications,* vol. 44, no. 12, pp. 43-49, 2012.

[30] P. Labbé, "Mobile Networks: the Effect of Relative Node Velocity on Effective Channel Capacity," in *(IEEE) 1999 International Conference on Telecommunications, ICT '99; Resource Management*, Cheju, Korea, 1999, pp. 147-156: IEEE, 1999.

[31] D. Chirca, "STANDARD TACTICAL DATA LINKS," *Land Forces Academy Review,* vol. 15, no. 3, p. 393, 2010.

[32] M. Dinc, "Design considerations for military data link architecture in enabling integration of intelligent unmanned air vehicles (UAVs) with navy units," in *NATO SCI Conference, Germany*, 2009, 2009.

[33] V. Jodalen, R. Otnes and B. Solberg, "Nato standards for HF communications: an overview and technical description," 2004.

[34] M. Jovanović and R. Todorović, "Joint tactical radio–aspects of standardization," *Scientific Technical Review,* vol. 53, no. 2, pp. 72-79, 2003.

[35] P. Labbé and R. Proulx, "Model-Based Measures for the Assessment of Engagement Opportunities Implementation and Test Results," Defence Research Establishment Valcartier (Québec), DREV - R - 9807, November 1998.

[36] R. Comroe and D. Costello, "ARQ schemes for data transmission in mobile radio systems," *IEEE Journal on*

Selected Areas in Communications, vol. 2, no. 4, pp. 472-481, 1984.

[37] P. Yu, X. Wang and H. Yu, "A novel ARQ scheme applied to wireless communication," in *2014 IEEE 5th International Conference on Software Engineering and Service Science*, 2014, pp. 1044-1047: IEEE, 2014.

[38] R. E. Hundt and L. Irving, "Final report of the public safety wireless advisory committee," *Report, PSWAC, September,* 1996.

[39] G. Bevacqua, J. Cacace, A. Finzi and V. Lippiello, "Mixed-initiative planning and execution for multiple drones in search and rescue missions," in *Twenty-Fifth International Conference on Automated Planning and Scheduling*, 2015, 2015.

[40] E. Weldon, "An improved selective-repeat ARQ strategy," *IEEE Transactions on Communications,* vol. 30, no. 3, pp. 480-486, 1982.

[41] E. Biglieri, M. Sciuva and V. Zingarelli, "Modulation and coding for mobile radio communications: channels with correlated Rice fading and Doppler frequency shift," *IEEE transactions on vehicular technology,* vol. 47, no. 1, pp. 133-141, 1998.

[42] E. Koski, "Concepts for a reliable multicast data link protocol for HF radio communications," in *MILCOM 2005-2005 IEEE Military Communications Conference*, 2005, pp. 681-687: IEEE, 2005.

[43] S. Medhekar, "Roll-Call: an energy efficient radio frequency identification system," Rutgers University-Graduate School-New Brunswick, 2007.

[44] D. Weissman, A. H. Levesque and R. A. Dean, "Interoperable wireless data," *IEEE Communications Magazine,* vol. 31, no. 2, pp. 68-77, 1993.

[45] W. C. Lee, *Mobile communications engineering*. McGraw-Hill Professional, 1982.

[46] A. M. Michelson and A. H. Levesque, "Error-control techniques for digital communication," *New York, Wiley-Interscience, 1985, 483 p.*, 1985.

[47] R. Sinha and S. Gupta, "Performance evaluation of a protocol for packet radio network in mobile computer communications," *IEEE transactions on vehicular technology,* vol. 33, no. 3, pp. 250-258, 1984.

[48] U. Svasti-Xuto, Q. Wang and V. Bhargava, "Capacity of an FH-SSMA system in different fading environments," *IEEE transactions on vehicular technology,* vol. 47, no. 1, pp. 75-83, 1998.

[49] S. L. Ariyavisitakul and G. M. Durant, "A broadband wireless packet technique based on coding, diversity, and equalization," *IEEE Communications Magazine,* vol. 36, no. 7, pp. 110-115, 1998.

[50] A. Fiser, "Mapping the long-term options for Canada's north: Telecommunications and broadband connectivity," 2013: Conference Board of Canada, 2013.

[51] M. Levitt, "Nation-building at home, vigilance beyond: Preparing for the coming decades in the arctic," Speaker of the House of Commons, Ottawa, 2019.