Video Testing at the FirstNet Innovation and Test Lab Using a Public Safety Dataset
Chris Budny, Jeffrey Liu, and Andrew Weinert

Abstract—Video applications are projected as a stressing and significant service of the Nationwide Public Safety Broadband Network. There is an need to empirically experiment and evaluate services across the actual network, rather than purely theoretical or simulation-based assessments. Our contribution to addressing this need, is being one of the first to demonstrate, using open source technology, how to conduct video experiments at the FirstNet Innovation and Test Lab. We transmitted video from the dataset across public safety and commercial broadband networks and under different network conditions. We demonstrated that data preemption and priority were dependent upon a variety of factors.

Index Terms—Cellular systems, communications, design methodology, long term evolution, mobile handsets.

I. INTRODUCTION

WITH the increasing frequency and cost associated with disasters, there is a critical need to develop technology to support incident and disaster response. The Nationwide Public Safety Broadband Network (NPSBN) is established and licensed by FirstNet and built and operated by AT&T for public safety. Video applications and analytics are routinely projected as a stressing and significant service of the NPSBN. However, there has been a formally identified dearth of datasets which are representative of, and tailored toward public safety operations to enable the development and testing of NPSBN capabilities optimized for public safety [1]. In response, based on outreach by Weinert and Budny [2] and informed by Palen et al. [3], a Public Safety Innovation Accelerator Program (PSIAP) video and imagery dataset of representative and operational public safety scenarios was developed by the New Jersey Office of Homeland Security and MIT Lincoln Laboratory (MIT LL). A key motivation for the development of this dataset was to enable NPSBN testing with public safety data.

A. Motivation

There are many efforts to simulate and evaluate long term evolution (LTE) performance [4], including previous modeling of the nationwide performance of the NPSBN [5]. Other research focuses on how to determine the efficient utilization of NPSBN bandwidth for video content transmission [6] or the Video Quality in Public Safety (VQiPS) working group[1] that helps public safety agencies with little or no technical expertise in video describe their video quality needs and provide basic guidance for the selection of key video system components. Majority of these efforts either focus on the underlying LTE network or the quality of experience of an video. With the potential to deploy unique public safety services on the NPSBN, there is an need to experiment and evaluate video services as a function of the LTE network, rather than continue to assess communications and applications independently.

B. Objectives and Contribution

Our objective was to be the first party external to FirstNet and AT&T to evaluate the effect of different LTE and NPSBN parameters on the quality of service (QoS) of streaming videos including the PSIAP dataset. We developed software to enable evaluation of NPSBN using the PSIAP dataset; in particular, the evaluation of video streaming performance. We do not contribute a full assessment of the NPSBN. This effort was also complimented by another effort to prototype a testbed for in-vehicle and edge computing testing [7].

II. EXPERIMENTAL DESIGN

We tested at the FirstNet Innovation and Test Lab in Boulder, CO over the week of May 20–24, 2019. A revocable license agreement between FirstNet and MIT LL was signed prior to testing. While not detailed in this letter, we worked with FirstNet to develop best practices and policies to enable others to test at FirstNet too.

A. Architecture

Fig. 1 provides the architecture diagram. A user equipment (UE) was placed inside an radio frequency (RF) isolation box and a virtual private network (VPN) connection was established between the UE and the FirstNet intranet which hosted the media server. A video file hosted on the UE would be streamed using the open source Ant Media Server2 implement real-time messaging protocol (RTMP) by an Android application. The application was designed for testing and not public safety deployment [8]. Previous research by

1https://www.nist.gov/ctl/pscrtools-video-quality-public-safety
2https://antmedia.io/
Aloman et al. [9] indicated that RTMP provided the best quality compared to other video streaming protocols. The video would be transmitted over the air (OTA) into the NPSBN where it would be ingested by the media server via an Internet connection. For UE downlink testing, the Open Broadcaster Software Studio\(^3\) was used to stream from the media server over the network to the UE.

To evaluate the performance of the network for streaming video, we compared the source and received stream videos to identify any quality degradations, such as dropped frames or visual artifacts. The UEs were two Samsung Galaxy S9+.

Furthermore, LTE multiple input, multiple output (MIMO) antennas are designed to be orthogonal and 90 degrees cross polarized. In a real world environment as in the RF isolation box, the angle of the device and its orientation to the base station antennas will impact device received power. We used the same UE orientation, as guided by tape, within the RF box to ensure a fixed received power across the repeating measurements. To minimize mechanical disruptions to the positioning of the UEs, we viewed and controlled the devices remotely over USB using the Vysor software.\(^4\)

### B. Use Cases

Given this architecture, we identified four potential network use cases, but scoped initial testing to just the first two use cases in Table I. Scope was limited to maximize testing over the most prevalent use cases. We also only focused on video streaming use cases and did not test the deployment of analytics, which could reduce the bandwidth required to communicate information. Analytic deployment on the UE or server was scoped as future work.

### C. Radio Access Network Loading

We used a radio access network (RAN) loader, managed by FirstNet, to simulate load on the network. We examined two types of simulated load: high-UE, and high-bandwidth.

![Architecture diagram.](image)

For the high-UE scenario, we simulated several hundred UEs connections to the network. Each simulated UEs remained in a data connected mode to the network through periodic ping requests. The RAN loading procedure has an initialization period, which took \(\sim 3\) mins to initialize. Once initialized, the RAN would linearly ramp the number of simulated UEs over 1 min. Once the network was fully loaded, the commercial UEs, but not the NPSBN UEs, would be disconnected from the network due to too many active connections. This allowed us to analyze preemption effects of commercial vs NPSBN UEs. For the high-bandwidth scenario, we simulated up to 50 concurrent commercial UEs downloading a 1 GB file from an FTP server. The procedure for this scenario involved a similar initialization and ramping phase like the previous scenario.

Due to time constraints, we did not simulate high-bandwidth uploads from UEs to a server. We also did not simulate the movement and velocity of the simulated UEs, although FirstNet has this capability.

### D. Video

We tested using samples of videos from the PSIAP dataset, which includes operational images and video from a variety of sources. Tested video was sourced from the Defense Visual Information Distribution Service (DVIDS), United States Geological Survey (USGS), Creative Commons video hosted on YouTube, and donated video from public safety and researchers. Across the tested videos, there was a wide range of scenarios, environments, lighting conditions, and obscurants. Fig. 2 provides a couple example video frames. We also split the videos into one minute clips using FFmpeg\(^5\) to standardize the video lengths between sources.

### III. Test Parameters

This section describes the five test parameters of attenuation, radio spectrum, QoS class identifier (QCI), allocation and retention priority (ARP), and subscriber identity module (SIM). We did not consider cell sector states of static, dynamic, or controlled, energy consumption by UEs, and other

---

3\(^{https://obsproject.com/}\)

4\(^{https://www.vysor.io/}\)

5\(^{https://ffmpeg.org/}\)
QoS parameters such as priority handling, packet delay budget, and packet error loss rate. We were not able to vary all of the parameters independently, as we only had access to a limited number of SIM card configurations.

A. Attenuation and Signal to Interference Noise Ratio

Attenuation is a reduction of signal strength (power) during transmission and is often referred to as signal loss. Attenuation can change due to many effects; one common cause is increased distance between the UE and the transmitter. We simulated various cell conditions using an RF attenuator: near-, mid-, and far-field conditions, corresponding to RF attenuation level of 0dB, 18dB, and 36dB. The measured signal quality attributes for each configuration are given in Tables II and III for the commercial and NPSBN SIM respectively.

The signal-to-inference noise ratio (SINR) is defined as the ratio of signal power to the noise and interference power, often expressed in decibels. Here the required SINR means that minimum level of SINR required to decode the LTE signal. The average LTE SINR is a function of the $S$ the averaged received signal power; $I$, the average interference power; and $N$, the noise power:

$$\text{SINR}_{\text{ave}} = \frac{S}{I + N}. \quad (1)$$

The Johnson–Nyquist noise (thermal) noise is approximately white, meaning that its power spectral density is nearly equal throughout the frequency spectrum. The Thermal noise at room temperature (15°C) is $-174 \text{ dBm/Hz}$. For Band 14 this would equate to $-104 \text{ dBm}$.

B. Radio Spectrum

Spectrum is the physical RF wavelength in which the signal is communicated over and is organized into ranges known as bands. We tested on Class 12 and 14. Class 12 and 14 consist of paired spectrum in the 700 MHz band. Class 12 consists of 698–716 MHz and 728–746 MHz; while Class 14 is allocated 758–768 MHz and 788–798 MHz with guard bands for interference mitigation at 768–769 MHz and 798–799 MHz [10]. For tests on the Class 12 network, we used the commercial production AT&T network. We limited the bandwidth used when testing over the production network as to not impact the public service. When testing with Class 14, we used both AT&T commercial and NPSBN SIM cards on the FirstNet lab network.

C. Quality of Service Class Identifier

The QCI affects the QoS performance characteristics. The current LTE standard defines nine QCI values as integers with each QCI assigned a priority level. In the case of “extreme” congestion the least important priority level traffic would be the first to be discarded. The QCI and priority level values do not always match.

We tested with two QCI values. First with a QCI of 6 and a priority level 6, which is assigned to FirstNet primary users (i.e., firefighters); and then with a QCI of 8 and a priority level 8, which is assigned to FirstNet extended primary users (i.e., animal control) or AT&T commercial users. Both these QCI values have non-guaranteed bit rate (GBR) resource types. We did not test with a QCI of 5, which has the highest priority per the 3GPP LTE standard.

D. Allocation and Retention Priority

ARP helps define pre-emption behavior across the network. Pre-emption is the network capability that permits authorized high priority traffic, such as FirstNet primary users, to take over resources assigned to lower priority traffic [12]. ARP consists of two Booleans of ARP-pre-emption vulnerability indicator (PVI) and ARP-pre-emption capability indicator (PCI) and a ARP priority level. We used the default ARP priority level and considered different values of ARP-PVI and ARP-PCI during testing. If ARP-PVI is true, then the UE is vulnerable to pre-emption and could be offloaded from the network. Conversely if ARP-PCI is true, then the UE is capable of pre-empting other UEs.

E. Subscriber Identity Module

The SIM directs the UE to which network to use. We controlled this by physically changing the SIM card in the UE in the RF box. We used SIMs for two networks: the AT&T commercial network and the AT&T NPSBN. The permitted QCI and ARP parameters were tied to specific SIM cards. Also, the AT&T NPSBN SIMs are prioritized over commercial users by the LTE scheduler. The three SIM cards that we used are listed in Table IV.
IV. Test Scenarios and Results

This section describes how the four test parameters described in Section III were organized to create test scenarios. Each test involved communicating video across the network, with each test lasting 2–3 minutes. We tested with the SIM cards listed in Tab. IV; as such, we were not able to vary all of the parameters independently, since the spectrum, QCI, and ARP configurations are determined by the specific SIM card. All test configurations were repeatable and could be replicated by future efforts.

We recorded the video stream at three different points along the UE uplink workflow: the raw unprocessed video that is fed into the stream, the output of the video stream after RTMP encoding on the uploading device; and the stream as transmitted over the LTE network on the receiving device. In total, we conducted one hundred tests across different network configurations, with some tests discarded due to various technical challenges.

A. Control Runs Without Simulated Load

We ran a set of control runs with no simulated load for the UEs, using each of the three SIM cards in Tab. IV. We first ran control runs for server to UE streaming, which consisted of streaming 1-minute clips from the PSIAP dataset to the UE from the media server. We documented network activity using the dstat linux utility on the server, as well as the diagnostic logs of the Ant Media Server; we also qualitatively verified the visual appearance of the received stream accurately reproduced the source stream. This was repeated for the UE to server upload streaming scenario as well.

These runs were performed at simulated near-, mid-, and far-field cell conditions using the RF attenuator. These control runs were done to ensure that video could be successfully transmitted both upstream and downstream to the UEs without additional simulated loads. This ensures that any disconnections that occurred during the test phase with simulated load were due to the simulated load and not environmental noise. All devices and SIM cards were able to complete the control trials of transmitting and receiving streaming video at all cell condition configurations. There was a small amount of visual distortion of the stream in the form of compression artifacts, as well as occasional drops in framerates. However, none were significant or frequent enough to meaningfully disrupt the quality of the received stream.

B. Test Runs With Simulated Load

We ran tests with both high-UE simulated load and high-bandwidth simulated load. For the high-UE scenario, we ran both uplink (UE to server) and downlink (server to UE) tests. We only considered SIMs configurations 2 and 3 for these tests, since we are interested in comparing the performance between the NPSBN and commercial AT&T configurations. SIMs 2 and 3 were each loaded into a UE, and placed in the same RF isolation box. Similarly to the control tests, dstat and server logs were collected to record network activity.

For the high-UE load uplink test, each UE loaded the same PSIAP test video and initiated a stream to the media server. The videos were played on a loop so that the streams would not end prior to the completion of the test. The high-UE RAN configuration was then run. Once the simulated load reached several hundred UEs, the device with the commercial SIM lost connection to the network, and the stream terminated. However, the device with the NPSBN SIM remained connected, and the stream continued. There were some dropped frames in the received stream, but not significantly more than in the control runs.

For the high-UE load downlink test, we streamed videos from the media server to both test UEs. Again, the videos were played in a loop, and the high-UE RAN configuration was started. Similarly, the device with the commercial SIM was disconnected from the network when the UE load exceeded approximately several hundred simulated devices; we observed choppy playback on the stream for a few seconds before the stream was disconnected. On the other hand, the device with the NPSBN SIM did not disconnect during this test, and the stream was maintained without significant disruption.

The results of the high-UE tests, both uplink and downlink, were consistent with the expected behaviors given the ARP values of the respective SIMs.

We then conducted a downlink test with high-bandwidth simulated load to model UEs using significant downlink bandwidth. This test was set up identically to the high-UE load downlink test, except we ran the high-bandwidth RAN configuration, which simulates UEs downloading large files from a server. We conducted five runs in this test corresponding to different numbers of simulated UEs: we first ran a test with 10 simulated concurrent UEs, each downloading a 1 GB file over File Transfer Protocol (FTP); we then increased the number of UEs by 10 for each subsequent test, up to 50 concurrent UEs. We observed and recorded the received streams on both the NPSBN SIM device, as well as the commercial SIM device. We found that both devices experienced comparable degradation in streaming video quality in this test. The streams on both devices exhibited significant packet loss, resulting in dropped frames and choppy playback; both streams also struggled to maintain the original video resolution. These effects were observed for 10 – 30 simulated UEs, with increasing load resulting in more dramatic degradation. At 40 and 50 simulated UEs, both streams were rendered unresponsive, and eventually lost connection to the server.

Due to other testing priorities and finite testing time available the week of May 20–24, we were unable to conduct an uplink test with high-bandwidth simulated load. We simply ran out of time. The results of the high-bandwidth tests illustrate a bottleneck in the testing architecture. Since the high-bandwidth load is transferred over the Internet, outside of the FirstNet network, the data priority policies of FirstNet do not apply. Thus, Internet traffic of NPSBN devices could be subject to congestive effects of a small number of commercial UEs if they are using significant bandwidth. This also highlights that the FirstNet network alone is not a comprehensive solution for prioritizing and managing public safety communications.
C. Challenges

We experienced issues with using a screen recording app simultaneously with uplink UE to server streaming. This configuration proved too computationally intensive for the UE, and resulted in low frame rates in streaming. We modified the testing application to instead save a copy of the encoded stream on the UE directly, instead of using an external application to record the stream contents after rendering. This allowed the UE to broadcast the stream at the original framerate.

We had also intended to test using an unlocked Samsung Galaxy Note 9 UE, model number SM-N960U1. However, we were unable to use the device with the NPSBN SIM card due to the unlocked model lacking the AT&T-specific features for IP Multimedia Subsystem/VoLTE over LTE (IMS/VoLTE) required to connect to the network. As such, we were unable to test using that device. Future tests should ensure that all UEs support the AT&T-specific features required for connection.

V. Future Work

Future work should further explore the effects of network resources and bandwidth usage on connection quality; additional PCI and PVI configurations; develop additional RAN loading configurations to explore different loading scenarios; and explore the utility of streaming output of analytics, which likely will require less bandwidth, rather than high bandwidth video. Evaluations should be for the downlink and uplink.

VI. Conclusion

We transmitted video included in the PSIAP dataset across the AT&T 4G LTE FirstNet NPSBN and commercial networks under different network conditions. We were the first party external to FirstNet and AT&T to demonstrate, using open source tools and data, how to conduct video experiments at the FirstNet Innovation and Test Lab. Due to the scope of the Boulder laboratory, we could not test other cellular networks in a controlled manner.

This initial and future experiment can explore how communication and information resources change depending upon public safety’s needs while informing which analytics or communication network management strategies need to be developed. We demonstrated that devices using FirstNet NPSBN SIM cards were not preempted due to a large number of UE connections, unlike devices using the commercial SIM. We also found that downlink transmission for both NPSBN and commercial devices were degraded by high-bandwidth usage from a small number of devices, due to the Internet traffic traveling outside of the FirstNet Core network, where data priority cannot be enforced.

Much of the software used is hosted on GitHub under BSD-2 licenses, managed by the MIT LL organization, https://github.com/mit-ll, with related repositories titled “PSIAP-*.” Details about how the PSIAP dataset was organized can be found in other peer review article [13] and a subset of the dataset is documented as the Low Altitude Disaster Imagery (LADI) dataset, https://github.com/LADI-Dataset.

ACKNOWLEDGMENT

The authors thank the National Institute of Standards and Technology, particularly John Garofolo and Scott McNichol. The New Jersey Office of Homeland Security and Preparedness have been excellent collaborators, including William Drew, Steve Talpas, Sean Gorman, Daniel Morroco, and Charles Agro. Lastly, without FirstNet support from Vihan Jani, Sirak Zewedie, and Michael P. Van Zuiden, this letter would not had been possible.

Approved for public release. Distribution is unlimited. (C) 2019 Massachusetts Institute of Technology. Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (Feb 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work.

REFERENCES

[1] T. McElvaney, R. Felts, and M. Leh, “Public safety analytics R&D roadmap,” NIST, Gaithersburg, MA, USA, Rep. 1917, Apr. 2016.
[2] A. Weinert and C. Budny, “Outreach to define a public safety communications model for broadband cellular video,” in Proc. IEEE Int. Symp. Technol. Homeland Security (HST), Oct. 2018, pp. 1–4.
[3] L. Palen et al., “A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters,” in Proc. ACM-BCS Visions Comput. Sci. Conf. (ACM-BCS), Edinburgh, U.K., Apr. 2010, pp. 1–12.
[4] A. Jarwan, A. Sabbah, M. Ibnkahla, and O. Issa, “LTE-based public safety networks: A survey,” IEEE Commun. Surveys Tuts., vol. 21, no. 2, pp. 1165–1187, 2nd Quart., 2019.
[5] R. Rouil, A. Izquierdo, M. Souryal, C. Gentile, D. Griffith, and N. Golmie, “Nationwide safety: Nationwide modeling for broadband network services,” IEEE Veh. Technol. Mag., vol. 8, no. 2, pp. 83–91, Jun. 2013.
[6] Y. Yang, H. Zheng, J. Chang, J. Contestabile, and S. Babin, “Advanced communication framework for video quality in public safety,” in Proc. IEEE Wireless Telecommun. Symp. (WTS), Chicago, IL, USA, Apr. 2017, pp. 1–6.
[7] L. Liu, X. Zhang, Q. Zhang, A. Weinert, Y. Wang, and W. Shi, “AutoVAPS: An IoT-enabled public safety service on vehicles,” in Proc. 4th Workshop Int. Sci. Smart City Oper. Platforms Eng. (SCOPE), Montreal, QC, Canada, Apr. 2019, pp. 41–47.
[8] P. Erickson et al., “Designing public safety mobile applications for disconnected, interrupted, and low bandwidth communication environments,” in Proc. IEEE Int. Conf. Technol. Homeland Security (HST), Waltham, MA, USA, Nov. 2013, pp. 790–796.
[9] A. Aloman, A. I. Ispas, P. Ciotirnae, R. Sanchez-Iborra, and M. D. Cano, “Performance evaluation of video streaming using MPEG DASH, RTSP, and RTMP in mobile networks,” in Proc. IEEE 8th IFIP Wireless Mobile Netw. Conf. (WMNC), Oct. 2015, pp. 144–151.
[10] J. C. Gallagher, “The first responder network (FirstNet) and next-generation communications for public safety: Issues for congress,” Congr. Res. Service, Washington, DC, USA, Rep. R45179, Apr. 2018.
[11] A. Kumbhar and I. Güvenç, “A comparative study of land mobile radio and LTE-based public safety communications,” in Proc. IEEE SoutheastCon, Fort Lauderdale, FL, USA, Apr. 2015, pp. 1–8.
[12] B. Kassa. Quality of Service, Priority, and Preemption. Accessed: May 2016. [Online]. Available: http://www.npsdc.org/download.jsp?tableId=37&column=217&id=3685&file=FirstNet_QPPIntro.pdf
[13] J. Liu, D. Strohschein, S. Samsi, and A. Weinert, “Large scale organization and inference of an imagery dataset for public safety,” in Proc. IEEE High Perform. Extreme Comput. Conf. (HPEC), Waltham, MA, USA, Sep. 2019, pp. 1–6.