

Wireless Streaming Video in California – Spring 2015

Ken Biba¹
Novarum, Inc.

CalSPEED collects sufficient raw data to analyze Internet video performance over cellular wireless networks. This analysis of wireless streaming and interactive video in California is based on the CalSPEED measurement survey of spring 2015.

- CalSPEED now analyzes both wireless streaming and two-way video conferencing services in California.
- More than 3/4ths of the sampled locations for AT&T, T-Mobile and Verizon are HD streaming-capable. Less than half of Sprint's locations are HD streaming-capable.
- Urban HD streaming served, on average, 38% more locations than rural HD streaming.
- Interactive video applications, like telemedicine and distance learning serve 50% more urban locations with HD service than rural locations.
- 50% fewer locations support interactive video to non-California locations compared to California locations. We speculate that this is due to differing quality of backhaul connections both within California and across the United States.
- Sprint's network has significantly fewer locations supporting either streaming or interactive video services compared to the other three carrier networks.

¹ Ken Biba is a consultant to the CPUC's CalSPEED project. He was tasked with preparing this analysis of the spring 2015 CalSPEED field test results. The content of, and conclusions reached in, this report are Mr. Biba's own, and do not necessarily reflect those of the State of California, the CPUC, its Commissioners or its employees.

1. Introduction

The CalSPEED program has been measuring wireless broadband performance and coverage of the four major wireless carriers since the spring of 2012. We now have seven rounds of statewide measurements spanning 3.5 years. These results include assessments of (TCP) data throughput, data latency and estimates of Over-The-Top packetized voice quality. For our Round 7 spring 2015 analysis, we have extended our scope to evaluate the ability of the mobile device to support one-way streaming video (such as YouTube) and two-way video conferencing.

We are using a modified measurement algorithm originated by Google to assess YouTube video quality and extend as our measurement metric. We examine two types of service – a streaming video service, such as YouTube, which is often delivered from a cache server located close to the user, and interactive video – a live two-way, video conference. Examples of this range from Skype, FaceTime, telemedicine and distance learning.

I would like to recognize my colleagues Dan Orr from California State University at Monterey Bay and Owen Rochte of the California Public Utilities Commission for their innovative work in creating maps of video quality.

2. Methodology

There is no consensus on an Over-The-Top streaming video metric, so we created one based on one that Google uses for YouTube². We then extended it for two way interactive video.

The Google YouTube quality metric is based on the three key quality measurements of the data stream:

²<https://www.google.com/get/videoqualityreport/#methodology>

- the Internet connection needs to work properly (no connection failures);
- the Internet connection must provide sufficient throughput to carry video of a certain quality; and
- the Internet connection must consistently maintain the throughput in order to smoothly deliver video without stuttering or pauses.

CalSPEED’s throughput tests make two ten-second throughput samples both upstream and downstream, from both a West Coast server (San Francisco Bay Area) and an East Coast server (Northern Virginia). Both locations are commonly used to host large streaming video storage and delivery servers and caches. Throughput measurements are taken every second, so we end up having twenty one-second throughput measurements in each direction, to each server.

If no connection can successfully be made for the test, we label the test as having no service³. We then assess the video service quality within each one second throughput sample according to a quality metric derived from the throughput necessary to support video quality levels.

Throughput (TCP) Criteria

| Sample Rating | Criteria | Reasoning |
|--------------------------|--------------------------|---|
| HD (High Definition) | >= 2.5 Mbps | Network throughput required for streaming HD (720p) video |
| SD (Standard Definition) | >= .7 Mbps < 2.5 Mbps | Network throughput required for streaming SD video (360p) |
| LD (Lower Definition) | < .7 Mbps | Network offers unreliable streaming video throughput |

³ We further filter tests by eliminating all test locations that are not within a carrier’s declared service area. We do not include measurements outside a given carrier’s declared service area.

| | | |
|-----------------|---------------|--|
| NS (No Service) | No connection | CalSPEED was unable to initiate a TCP connection from the user device to the CalSPEED measurement server |
|-----------------|---------------|--|

With each one-second sample measured, we then estimate the quality and consistency of the entire 20-second stream by the percentage of the 20 samples that equal or exceed the quality metric. Following the standard adopted by Google, we use a threshold of 90% of samples at or above the quality metric to assign a quality metric to the entire stream.

Sample Video Quality Criteria

| Stream Rating | Criteria | Reasoning |
|--------------------------|-------------------------------------|--|
| HD (High Definition) | 90+% samples are marked HD | Network throughput required for streaming HD (720p) video |
| SD (Standard Definition) | 90+% samples are marked at least SD | Network throughput required for streaming SD video (360p) |
| LD (Lower Definition) | Neither of the Above | Network offers unreliable streaming video throughput |
| NS (No Service) | No connection | CalSPEED was unable to initiate a TCP connection from the user device to the CalSPEED measurement server |

For streaming video, we assume that video is locally cached at the West server. We use the West unidirectional downstream video stream quality metric to assign a stream video metric to each location for each carrier. Quality varies not only based on network quality but also on the quality of the user device. Android smartphones are the consistent device used over time, so the reported results are limited to smartphones.

Unlike video streaming, video conferencing⁴ uses bidirectional video and audio either to a local user (here, modeled by the West server) as well as a distant user anywhere else on the Internet (here modeled by the East server). We model an interactive video session by using both the upstream and downstream video quality metrics combined with the OTT audio MOS metric.

We measure the interactive video quality as follows:

| Stream Rating | Criteria | Reasoning |
|--------------------------|---|--|
| HD (High Definition) | Both upstream and downstream video streams are HD and the MOS metric is ≥ 4.0 | Network offers consistent and reliable bidirectional interactive video HD (720p) and audio performance |
| SD (Standard Definition) | Both upstream and downstream video streams are at least SD and the MOS metric is ≥ 4.0 | Network offers consistent and reliable bidirectional interactive video SD (360p) and audio performance |
| LD (Lower Definition) | Either upstream or downstream video streams are not at least SD or the MOS < 4.0 | Network offers unreliable interactive video and audio performance |

⁴We should note that the interactive analysis is likely optimistic, since one end of our modeled connection is always a CalSPEED measurement servers in major urban areas while the other side of these connections can be user devices either in urban or rural locations. We do not model devices directly connected to other user wireless devices.

| | | |
|-----------------|---------------|--|
| NS (No Service) | No connection | CalSPEED was unable to initiate a TCP connection from the user device to the CalSPEED measurement server |
|-----------------|---------------|--|

Now let's look in Section 3 at what these metrics tell us about streaming video quality in the spring of 2015 in California. In Section 4, we will then look at video conferencing quality.

3. Streaming Video Quality

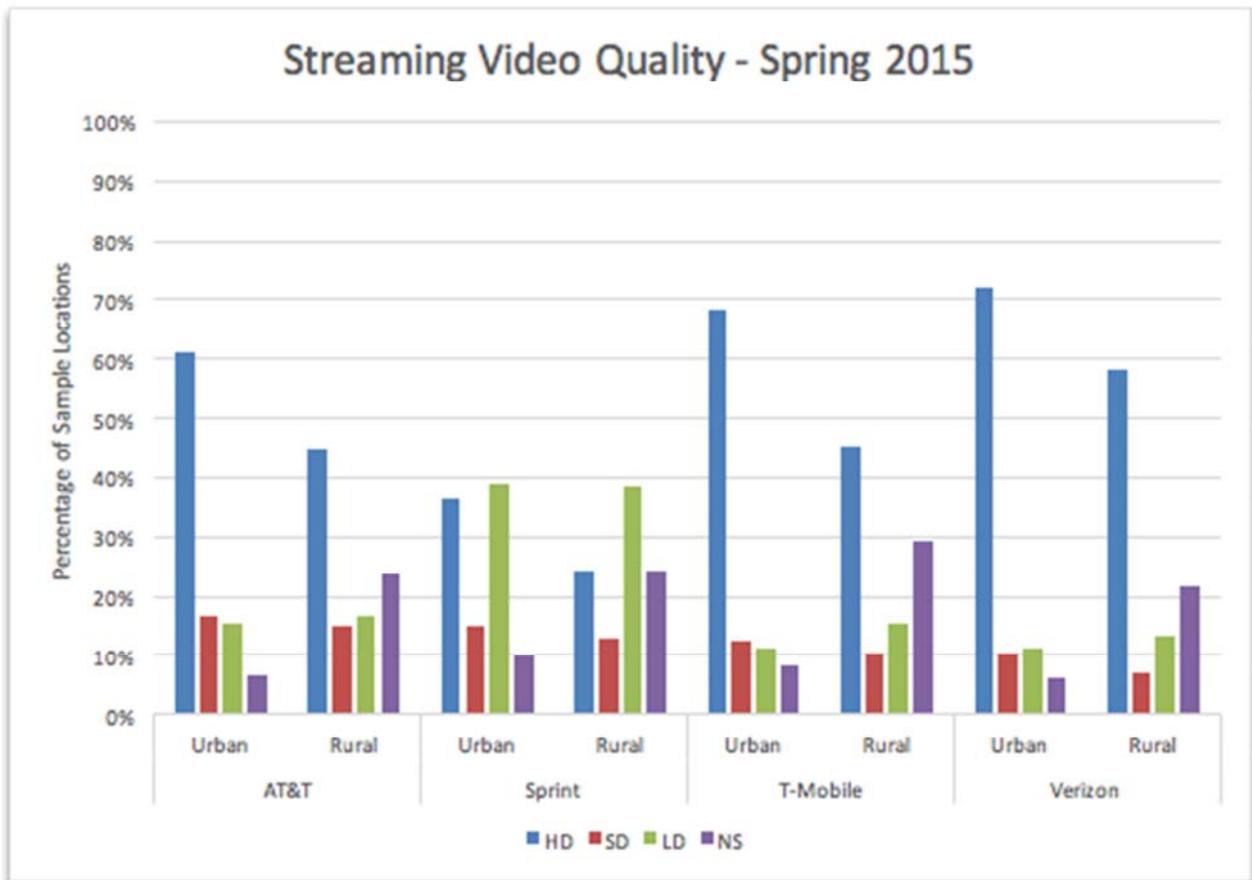
The streaming video metric is based on the downstream throughput quality from the West server, which models the nearby Internet caching of video content⁵.

A number of patterns are apparent.

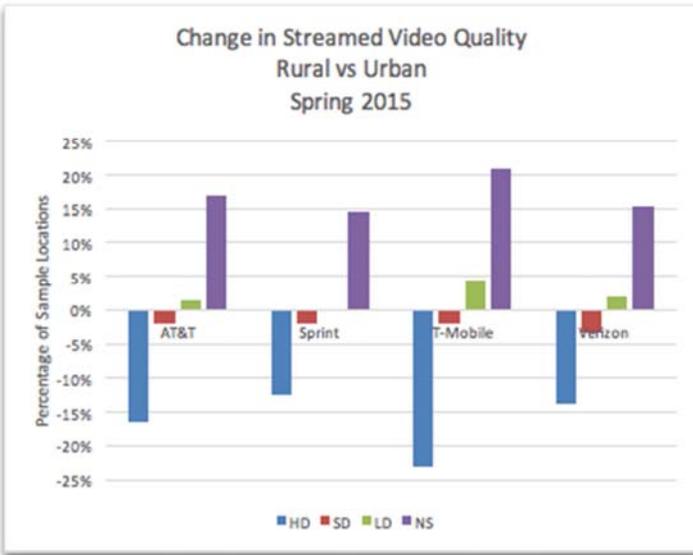
1. The quality of wireless streaming video quality differs materially between carriers.
2. HD quality streaming video is supported at more than half the locations of AT&T, T-Mobile and Verizon.
3. There is a material difference between the number of locations supporting HD compared to SD quality streaming video. There are relatively few locations that only have SD video. Essentially if a location supports streaming video, then it is almost always HD quality. For AT&T, T-Mobile and Verizon, the percentage of locations supporting either SD or LD service is constant at ~20% for urban and rural.
4. There is a substantial degradation of service when one moves from urban to rural areas. This degradation is not only about a decrease in video quality, but also a dramatic increase in locations with no video capability at all. For AT&T, we can see the number of locations with HD quality drop from 60% in urban to 45% in rural while the number of locations with no service more than quadruples from 5% to 23% in

⁵ The results are percentages of tested locations filtered by the asserted coverage areas for each carrier.

rural areas. For Sprint, the number of locations with HD availability drops from 37% in urban to 23% in rural, while the number of locations with no service more than doubles from 10% to 23%. For T-Mobile, the number of locations with HD video service drops from 87% to 45% in rural areas, while the number of locations with no service increases from 7% in urban areas to 29% in rural areas. The number of locations with Verizon HD video service drops from 72% in urban to 67% in rural while the number of locations with no service rises from 4% in urban areas to 22% in rural areas.

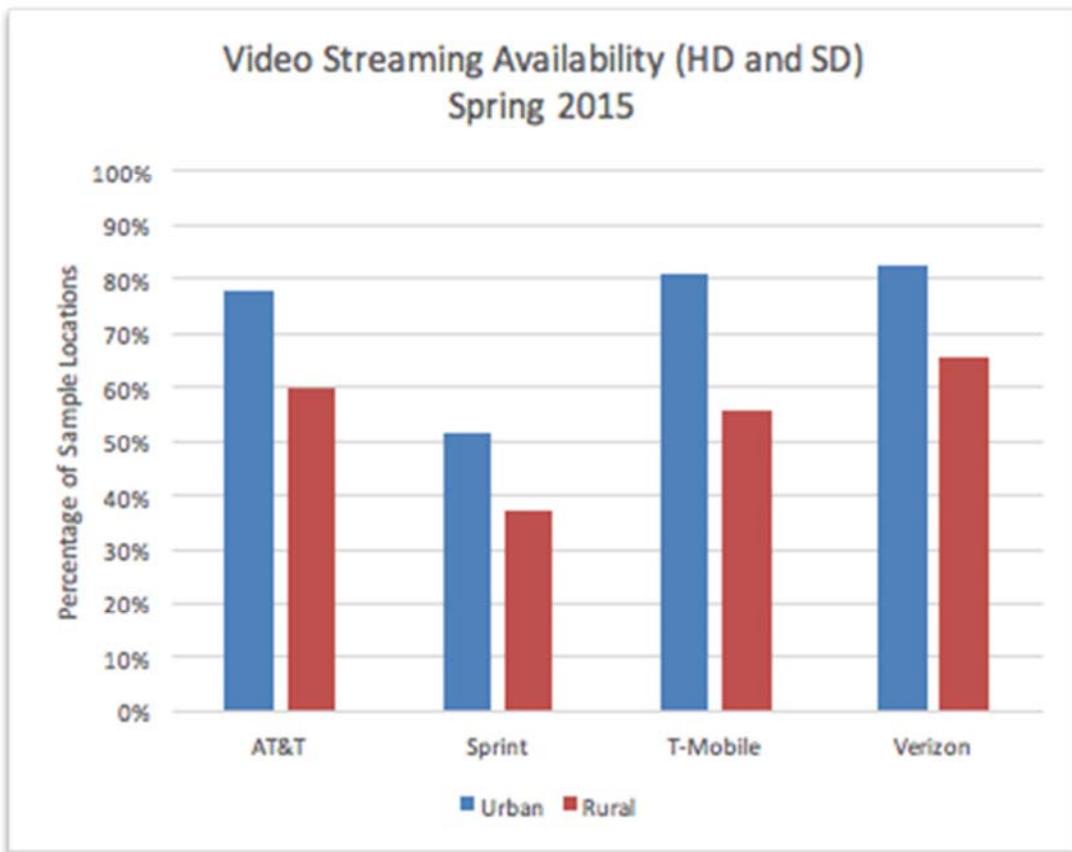


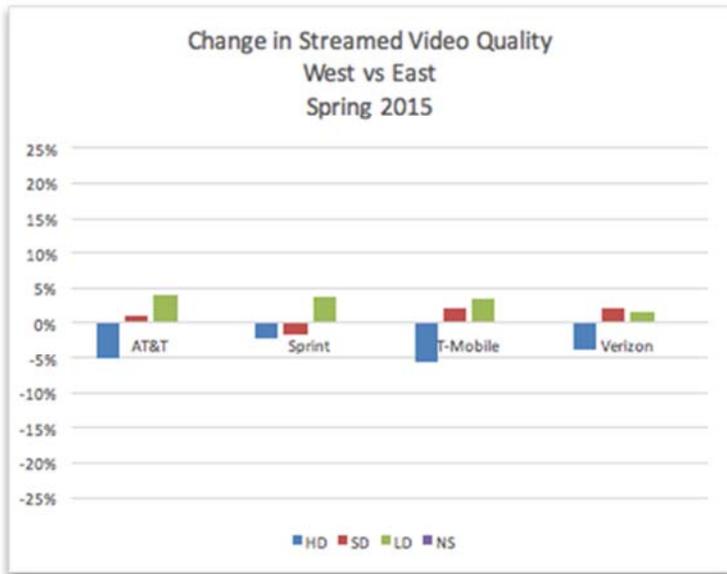
Let's look a bit closer at what happens to rural users of streamed video as compared to urban users.



As we move from urban to rural we see modest decreases in SD service ... but dramatic decreases in HD service accompanied by dramatic increases in no service. Rural users get dramatically less access to streaming video than do urban users.

If we define video streaming availability as the presence of either HD or SD service, we can see a consistent pattern of decreasing streaming video availability across all carriers as we from urban areas to rural - on the order of 20%.



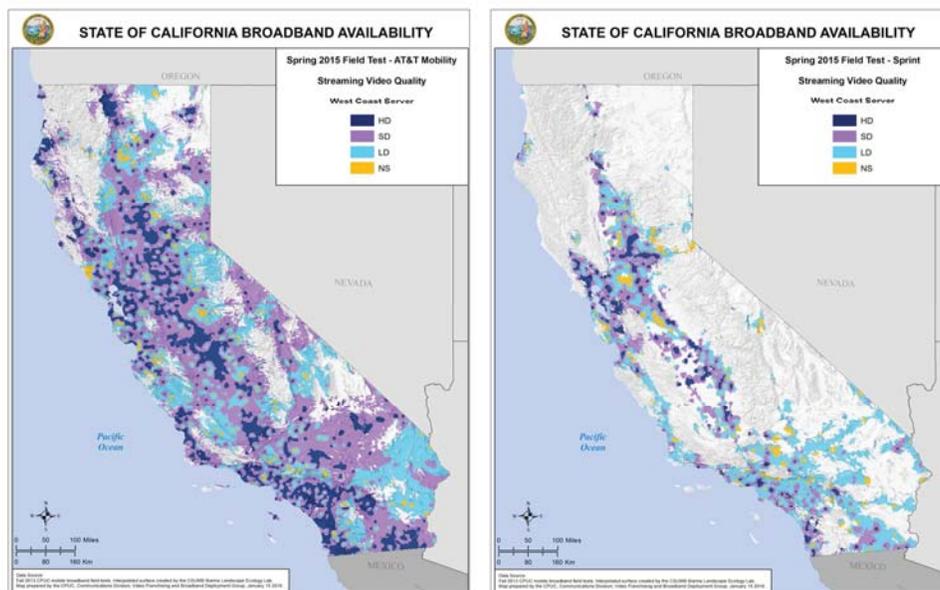


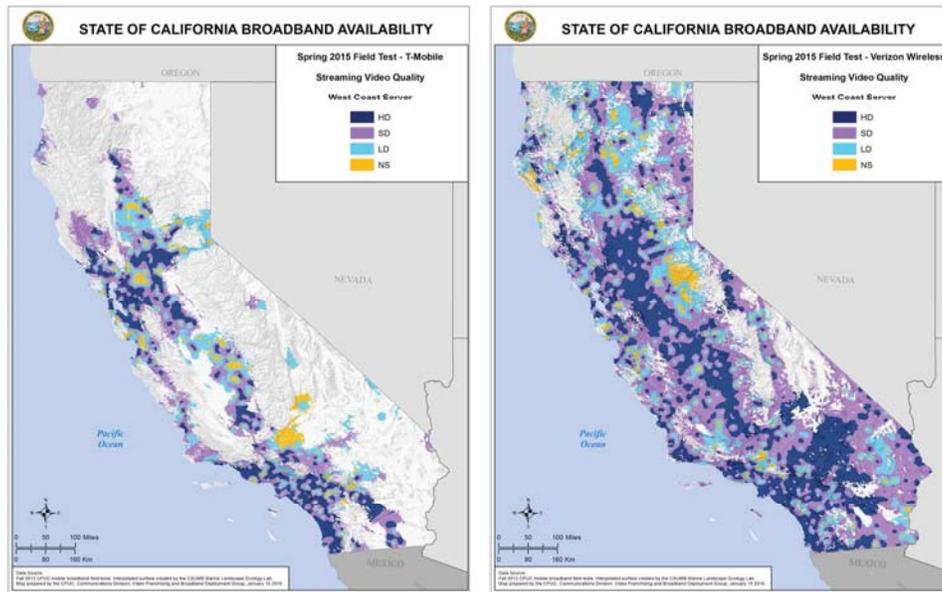
It is interesting to also look at the streaming quality from the East server. There is only modest change from the quality of the West server, with no change in no service and a modest decrease in HD quality and an increase in SD and (particularly) LD quality.

We have also made maps of streaming video quality in California using an extension of our existing network quality mapping methodology (see Appendix A).

The maps illustrate several points rather dramatically.

1. There are material differences between mobile streaming video among the four carriers both in coverage and in quality.
2. HD quality streaming video is concentrated in urban areas and along major transportation routes.





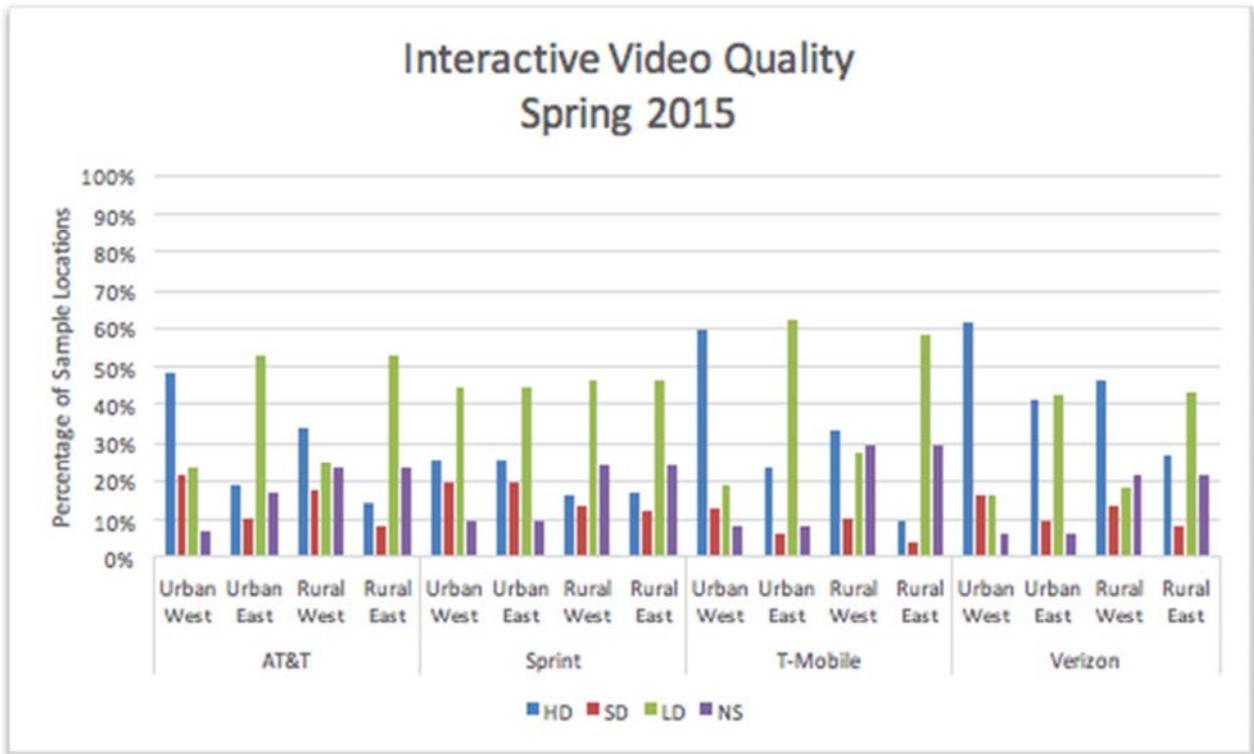
4. Two-Way Video Conferencing Quality

We look at three components that can be used to measure two-way video conferencing quality – the downstream video quality, the upstream video quality and the voice quality (as determined by the MOS score). For this analysis, we are using the measurement servers as proxies for another user – likely using a lower quality connection than the server uses. As such, this analysis should be considered an optimistic estimate. As we shall see, there is a large difference between urban and rural service - and one destination in our measurement is always urban - because it is the server location.

Two-way conferencing is a tougher standard than streaming - it includes both directions of video streaming (and typically upstream throughput is materially lower than downstream) as well as the tough standard of OTT⁶ MOS⁷ for real-time voice.

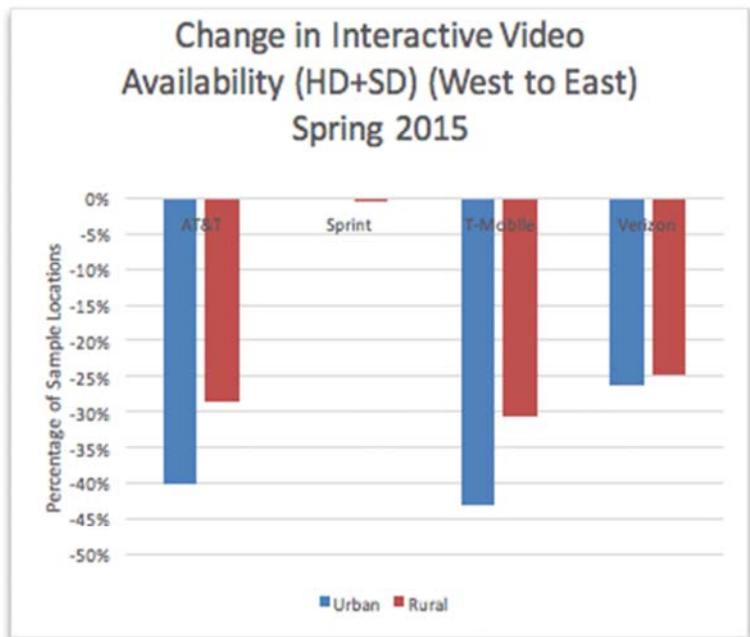
⁶ Over-The-Top. Streaming video on top of TCP.

⁷ We require an OTT MOS score of 4.0 or greater for both SD and HD.



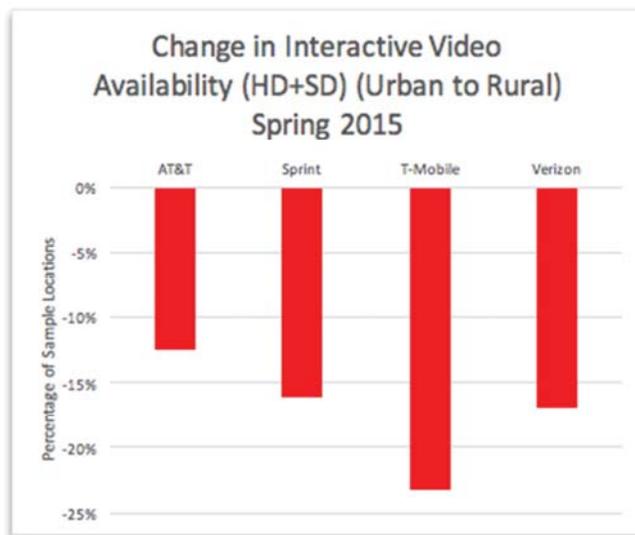
The above chart above illustrates several unique aspects of two-way video conferencing.

First, there is dramatic difference in interactive video quality for AT&T, T-Mobile and Verizon when the server side of the session moves from the West coast to the East coast. This is true for both urban and rural users. This move simulates a two-way conference session to a user across the country. This suggests that carrier Internet backhaul design for these three carriers substantially reduces interactive video quality. The number of locations supporting either HD or SD quality decreases as distance to the server



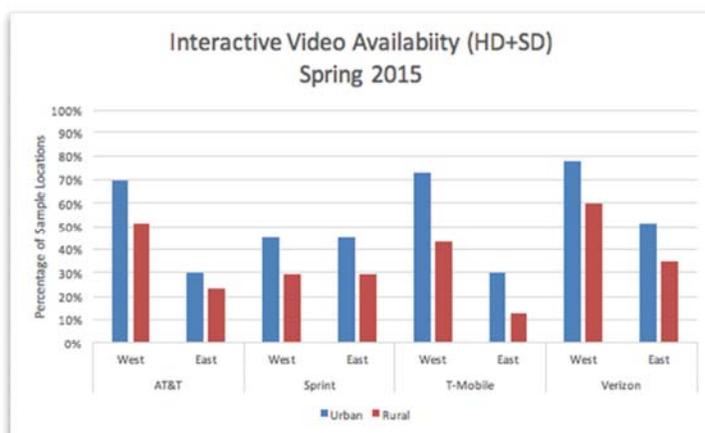
increases, while the number of locations with LD quality dramatically increases. Users would likely experience this not only as an increase in session failure, but also as an increase in “successful” sessions of unusable quality. Sprint shows very little difference in video quality between the two servers - generally poor for both West and East as well as for urban and rural locations.

Second, there is a substantial (though less so than what is caused by Internet distance to the server) quality difference in two-way video conference quality between urban and rural users. The chart to the right illustrates that the percentage of locations supporting either HD or SD interactive video session decreases by 10-20% for all carriers as the model user moves from urban to rural locations.



These effects are additive - so a rural user attempting a two-way video session would be less likely to have HD or SD quality, and even less likely if the distance to the called party were large. For example, for AT&T the percentage of HD interactive sessions decreases by two thirds from 48% availability for an urban user to the West server, to 13% for a rural user to the East server. And for Verizon, the percentage of HD-supporting locations decreases by more than half from 61% for an urban user to the West server, to 27% for a rural user to the East server. For T-Mobile, the effect is even more pronounced decreasing from 60% to 10%!

We can summarize this analysis on the chart to the right. It illustrates the total percentage of sample locations that can support either HD or SD quality interactive video sessions. We can see both



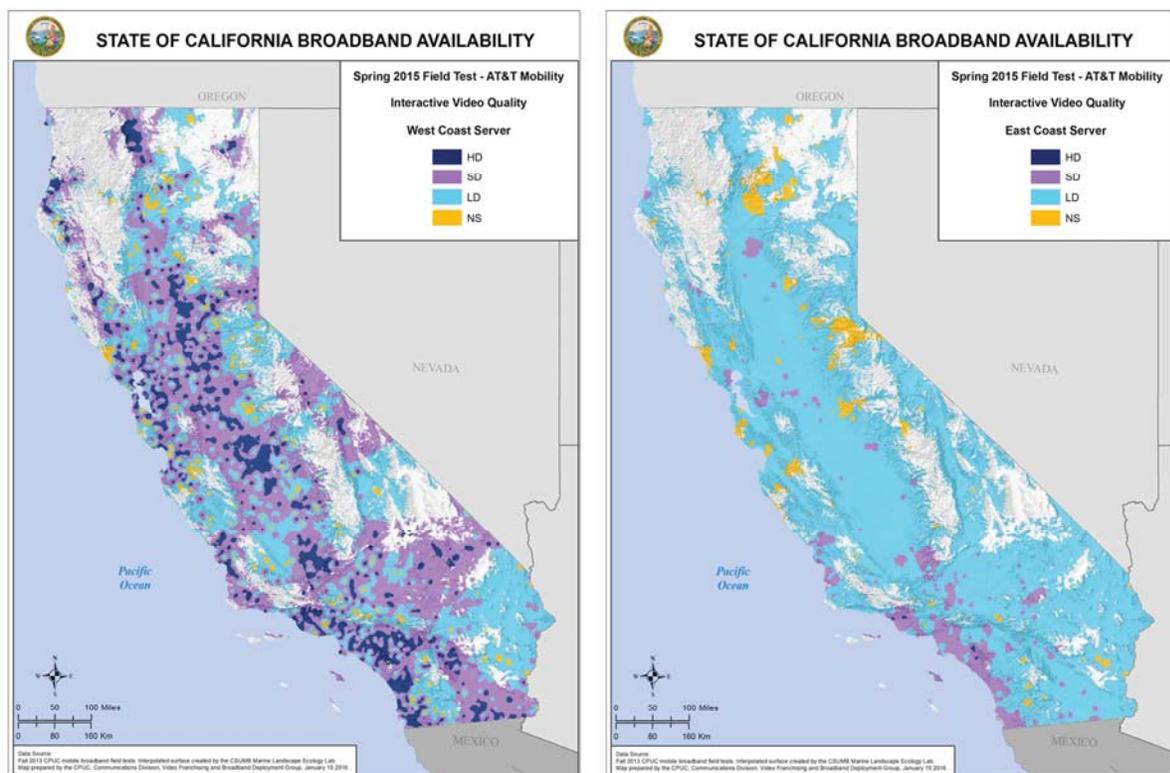
the quality degrading effects of Internet distance (between West and East servers) and between urban and rural users.

These effects can best be seen in maps of two-way conferencing quality in California.

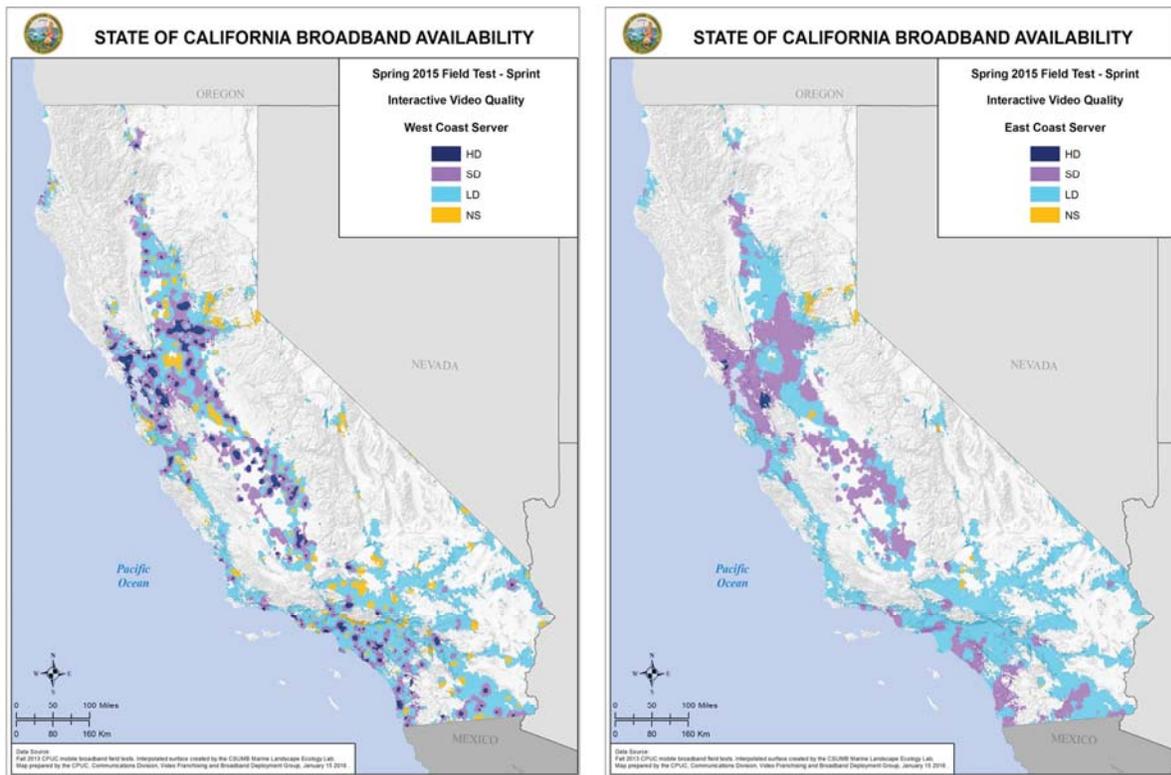
For AT&T, the map on the left illustrates two-way conferencing quality where the called party is simulated by the West server. We can see that HD and SD quality sessions are largely limited to urban areas and transportation corridors.

The map on the right illustrates two-way conferencing quality where the called party is simulated by the East server. The decrease in quality is marked. Very few locations in California, either urban or rural, will deliver a HD or SD quality two-way conferencing to a modeled urban user on the East coast.

Since the only difference between these scenarios is the Internet backhaul used by AT&T to get to the East server, these routes matter.

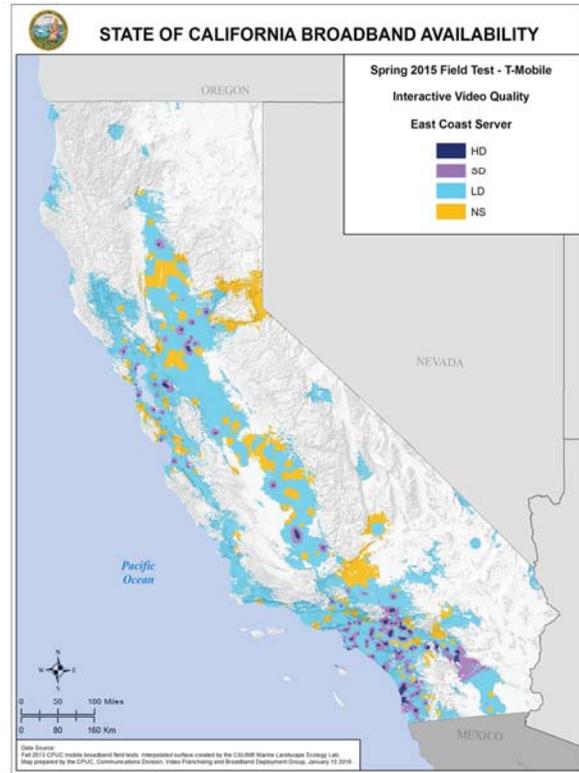
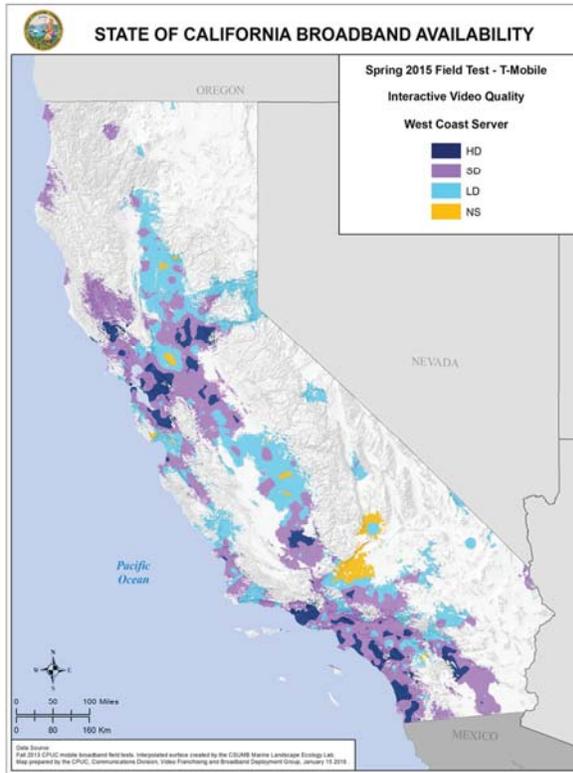


Now let's look at Sprint.

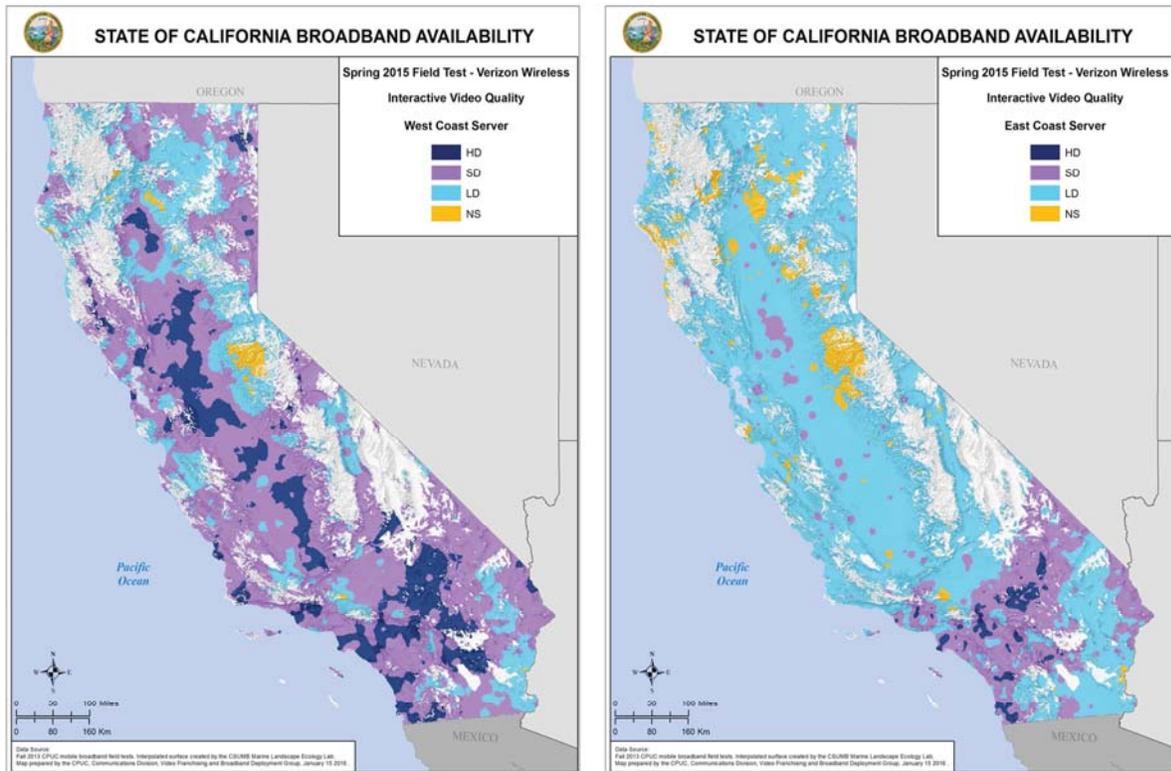


In the map at the left, we can again see the difference between being an urban user vs. a rural user - a decrease in quality for rural users. But it is the map on the right that is more interesting. Unlike AT&T, there is a less marked decrease in the extent of quality degradation when in session to the East coast. Sprint has made different backhaul choices that result in less difference about location.

T-Mobile exhibits the same trend for decreased service quality to rural users but introduces a slightly different effect of Internet distance. Notice the clusters of higher quality service in Southern California. This is suggestive that the T-Mobile Internet backhaul that T-Mobile uses in Southern California might provide slightly better service to two-way conference sessions to the East coast.



The maps for Verizon demonstrate both effects vividly. The map on the left dramatically shows much more extensive HD quality service in urban areas and transportation routes than in rural areas.



The map on right, for Verizon, is very interesting. Note the effect for most of California dropping from HD/SD quality to LD quality - except for a very specific band in Southern California. This specific area of California would appear to have much higher two-conferencing quality to the East Coast than any other area of California for any carrier! This suggests a better backhaul capability for Verizon in this area of California than in other areas of the state.

5. Conclusions

- CalSPEED now analyzes both wireless streaming and two-way video conferencing services in California.
- More than 3/4ths of the sampled locations for AT&T, T-Mobile and Verizon are HD streaming-capable. Less than half of Sprint's locations are HD streaming-capable.
- Urban HD streaming served, on average, 38% more locations than rural HD streaming.

- Interactive video applications, like telemedicine and distance learning, serve 50% more urban locations with HD service than rural locations.
- 50% fewer locations support interactive video to non-California locations compared to California locations. We speculate that this is due to differing quality of backhaul connections both within California and across the United States.
- Sprint's network has significantly fewer locations supporting either streaming or interactive video services compared to the other three carrier networks.

Appendix A: Video Mapping Analysis Model Validation

7th Round CalSpeed Video Analysis Model Validation

Here we used a Cohen's Kappa analysis to compare measured values versus model predicted values. Cohen's Kappa is a statistic used to compare agreement between categorical variables such as the ones we are using here. We used the numerical ranking in the comparison:

HD =3, SD = 2, LD = 1, NS = 0

The Kappa statistic describes agreement where 1 = complete agreement and 0 = no agreement. Cohen's Kappa is more statistically sound than a percent agreement calculation because it takes into account the agreement occurring by random chance. For example if you have a Cohen's Kappa of 0.79 you would say that the two variables of interest (here measured values and model predicted values) are in 79% agreement when you take away percent agreed upon by random chance.

The specific method we used was a quadratic weighted Cohen's Kappa. This was done because the data is ordinal (ranked in an order such as high to low and not just random categorical values).

This validation analysis accounts for the categorical nature of the variable of interest as well as its ordinal nature. It measures the agreement between the measured category and model predicted category, while taking into account the amount of agreement occurring by random chance. This analysis was done in R 3.2.2 using the IRR package.

| Provider | Service Type | Kappa |
|----------|-----------------------|-------|
| AT&T | Video Conference East | 0.718 |
| AT&T | Video Conference West | 0.997 |
| AT&T | Video Streaming West | 0.996 |
| Sprint | Video Conference East | 0.544 |
| Sprint | Video Conference West | 0.997 |
| Sprint | Video Streaming West | 0.983 |
| T-Mobile | Video Conference East | 0.997 |
| T-Mobile | Video Conference West | 0.744 |
| T-Mobile | Video Streaming West | 0.997 |
| Verizon | Video Conference East | 0.796 |
| Verizon | Video Conference West | 0.614 |
| Verizon | Video Streaming West | 0.998 |

These values will allow for interpolation of categorical values. Interpolation assumes that the data is continuous and so the resulting interpolations are continuous. Raw raster models were processed so that values were rounded up if it was 0.5 or more above the preceding integer and values were capped at 3. This resulted in a return to categorical values that could then be converted to shapefiles to include categorical labels of HD, SD, LD and NS.