Upstream OFDMA Anomaly Detection and Triaging

A Technical Paper prepared for SCTE by

Jay Zhu
Senior Engineer
CableLabs
j.zhu@cablelabs.com

Karthik Sundaresan
Distinguished Technologist
CableLabs
k.sundaresan@cablelabs.com

CableLabs
858 Coal Creek Circle, Louisville, CO,80027
3036619100

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

1. Introduction

Upstream Orthogonal Frequency Division Multiple Access (OFDMA) technology in DOCSIS 3.1 is starting to be rolled out in the field. Operators are beginning to test Upstream OFDMA channels in the lab and in the field and are discovering various intricacies in getting the upstream OFDMA to work robustly. Lower frequencies in the upstream spectrum can be noisy and making use of those portions of the spectrum tougher. Upstream RxMER looks very different than the Downstream RxMER, due to the noise funneling characteristics on the HFC plant, the additive nature of noise has a large impact at the CMTS upstream receiver.

As operators roll out OFDMA technology, they are starting to collect data on the performance of these OFDMA channels. This includes the US RxMER data, IUC usage hours, profile definitions etc. As a cable industry we are just starting to comprehend the OFDMA channel performance. Analyzing the US RxMER data and the IUC data is a powerful tool in understanding the performance of each of the node segments and the individual modems. This paper will discuss methods on how to analyze the upstream network data. It will discuss algorithms on how to logically extract the outlier modems and node segments. This paper will discuss methods for anomaly detection, historical behavior analysis, pattern recognition, classification and condition evaluation in the access network data. Combining the analysis of data along, with network topology and device location, it is possible to create a general view of the plant condition and isolate problem sources. The paper will implement methods on how to assign a health score to modems and network segments in an effort to triage which are the top priority nodes that operators need to work on. All this will enable operators to reduce upstream OFDMA troubleshooting and problem resolution time, reducing operational costs and enhancing network reliability.

2. Upstream Background

DOCSIS® 3.1 is now largely deployed in the field. This has primarily focused on a very successful roll out of the Downstream Orthogonal Frequency Division Multiplexing (OFDM) technology. Operators now are beginning to test Upstream Orthogonal Frequency Division Multiple Access (OFDMA) and are now deep into understanding the various intricacies in getting the US OFDMA to perform robustly. The first step is for an operator to get a good stable initial configuration of the OFDMA channel (location channel parameters, IUC definitions etc.). Once an operator can get CMs operating reliably on the OFDMA upstream channel, the operator can then start thinking about how to improve the reliability and efficiency of that upstream channel.

In the upstream direction, the cable system may have a 5-42 MHz, 5-65 MHz (Europe), 5-85 MHz, or 5-204 MHz pass bands. While a DOCSIS 3.1 CM supports a minimum of two independently configurable OFDMA upstream channels with each occupying a spectrum of up to 95 MHz, the challenge has been to find appropriate space in the spectrum to locate these channels. Operators who are running a 5-42 MHz plant are trialing out OFDMA in the space available after the spectrum used by 3 or 4 SC-QAM channels. Operators with mid split (5-85) plants usually have up to 10 SC-QAM channels and are making space for an OFDMA channel by using some of the spectrum just below 85 MHz and turning off a few SC-QAM channels. Operators in Europe with a 5-65 plant typically have 3 or 4 SC-QAM channels are using the remaining space an OFDMA channel, either the OFDMA channels go from 20-45 MHz or 45-65 MHz, depending on where the SC-QAM channels are.

The OFDMA upstream multicarrier system is composed of either 25 kHz or 50 kHz wide subcarriers. For a 95 MHz channel, this equals 3800 25 kHz spaced subcarriers or 1900 50 kHz spaced subcarriers. DOCSIS 3.1 Upstream transmission uses OFDMA frames. Each OFDMA frame is comprised of a configurable number of symbols (K = 6 to 36). Several transmitters may share the same OFDMA frame by transmitting on allocated subcarriers of the OFDMA frame. The upstream spectrum is divided into groups of subcarriers called minislots. Minislots have dedicated subcarriers, all with the same modulation order ('bit loading'). OFDMA minislots are 400 kHz wide and have either 8 (50 kHz) or 16 (25kHz) subcarriers. The modulation order of a minislot, as well as the pilot pattern used may change between different transmission bursts and are determined by the profile definition.

2.1. Upstream Network Data

CMTS and CM support features and capabilities that ca be leveraged to enable measurement and reporting of network conditions. These Proactive Network Maintenance (PNM) features deliver metrics which operators can use to identify undesired impacts such as plant equipment and cable faults, interference from other systems and ingress noise. With this information operators can make modifications necessary to improve conditions and monitor network trends to detect when network improvements are needed.

The OFDMA technology comes with a few different PNM measurements: Upstream Capture for Active and Quiet Probe, Upstream Triggered Spectrum Analysis, Upstream FEC Statistics, Upstream RxMER Per Subcarrier, Upstream Equalizer Coefficients, Upstream Impulse Noise Statistics, Upstream Histogram, Upstream Channel Power. These eight features are detailed in the [PHYv3.1] specification. The purpose of these upstream PNM functions is to analyze the upstream in various ways. These goals include measuring the plant response, understanding the underlying noise floor, having a wideband spectrum analyzer function on the CMTS, gathering statistics of burst/impulse noise occurring in a selected band, understand the linear response of the upstream cable plant, monitoring upstream link quality via FEC and related statistics, understanding the nonlinear effects in the channel such as amplifier compression and laser clipping, providing an estimate of the total received power in a channel and becoming aware of the upstream receive modulation error ratio (RxMER) for each subcarrier.

2.1.1. US RxMER

A CMTS uses upstream probes for ranging-related functions such as determining transmit pre-equalizer coefficients. A CMTS also uses the upstream probe to take an RxMER (received modulation error ratio) measurement. The CMTS grants probe opportunities to a CM in a P-MAP message with the "MER" bit set. When the CMTS receives the probe transmissions from the CM corresponding to such a grant, it performs the RxMER measurement and uses the results in its decision making. It also populates the corresponding MIB object or can upload a RxMER per subcarrier file via TFTP, for the operator's information. Some CMTS implementations also measured the RxMER in an alternate fashion, they are using the actual data transmission bursts from a CM to do a measurement.

2.1.2. Other US PNM data

The other Upstream PNM features (Upstream Capture for Active and Quiet Probe, Upstream Triggered Spectrum Analysis, Upstream FEC Statistics, Upstream Equalizer Coefficients, Upstream Impulse Noise Statistics, Upstream Histogram, Upstream Channel Power) are in various stages of maturity on different CMTS platforms.

As of the writing of this paper, we have had access to a lot of Upstream RxMER data, but have not been able to gather a meaningful set of samples from the field for any of the other PNM data types. In the future, as we get access to more data samples of those types from CMTS in the field, we can start analyzing those (for a future paper). For this paper, we focus on the US RxMER data that we have and start building methods and tools and ways to visualize and analyze this data set and figure out what kinds of upstream evaluation we can perform.

3. Upstream Observations

3.1. Introduction to Upstream RxMER Data

The upstream RxMER data is a PNM metric measured by the CMTS to report receive modulation error ratio on a per subcarrier basis. Similar to the downstream RxMER data, the upstream RxMER data can be helpful in identifying upstream impairments over frequency as well as performance fluctuations over time. The following is an upstream RxMER capture measured from 23.9 MHz to 51.4 MHz. Typically, noise floor is higher at lower frequencies. Although Pre-Equalization can help compensate the tilt, we still observe a slight inclination in the data from the lower frequencies to the higher frequencies.

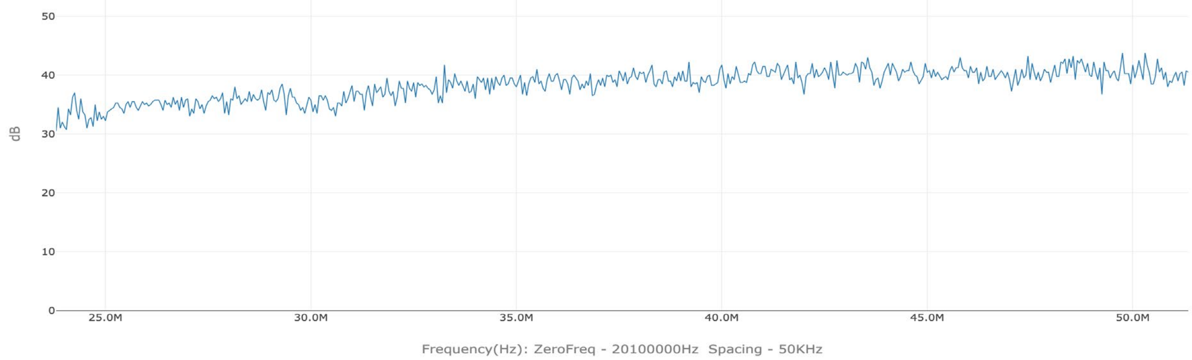


Figure 1 – An US RxMER sample on a OFDMA channel from Single CM

When more samples are captured over time and are displayed in one plot, we can observe the changes of RxMER values over time. By capturing more upstream RxMER samples and analyzing the over-time features of the data, there is opportunity that intermittent impairments and field events can be observed and categorized. In this paper, the discussion is focused an upstream RxMER dataset that was collected from multiple OFDMA interfaces during a 2-week timeframe.

Figure 2 – Multiple US RxMER samples, OFDMA channel, Single CM (lab)

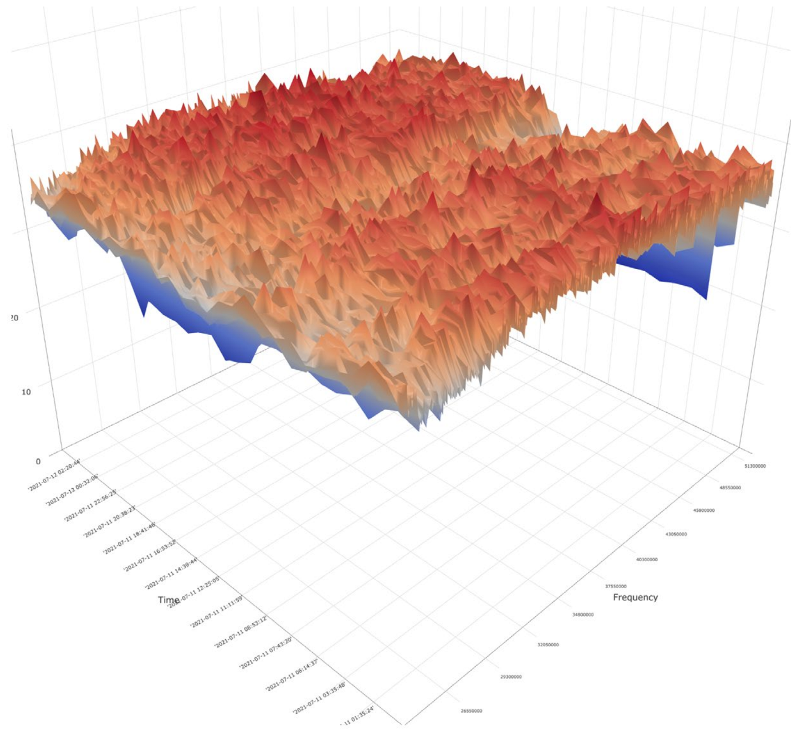


Figure 3 – 3D view of US RxMER samples, time versus frequency

3.2. Variations and Impairments

The following figure shows upstream OFDMA RxMER samples collected from a CMTS in the field. A lot of variation in RxMER values is observed in this visualization. The RxMER values of unused/inactive subcarriers are reported by the CMTS as 0xFF (63.75 dB), which are the spikes seen in the graph. The maximum difference between the RxMER values on a same subcarrier can be as much as 20 dB, which indicates that the condition of the upstream OFDMA channel is constantly changing. Ingress noise can also be observed on captured samples.

Figure 4 – Multiple US RxMER samples, OFDMA channel, Single CM (field)

In the RxMER captures collected from a different CM, we observe less variation in RxMER values. We also observe standing waves across the spectrum.

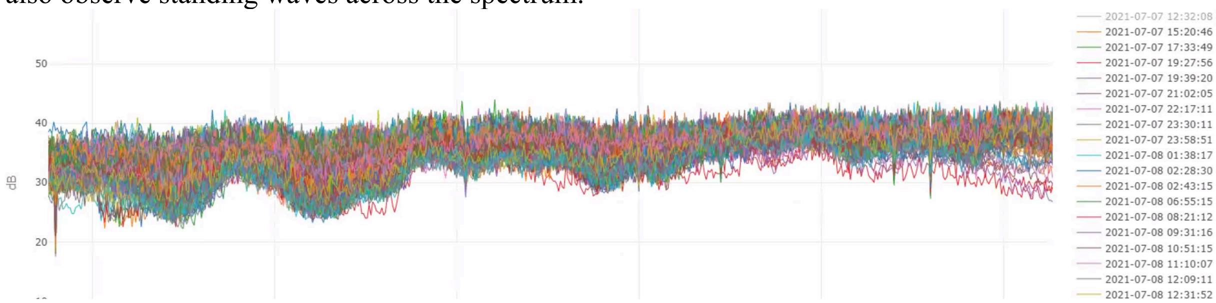
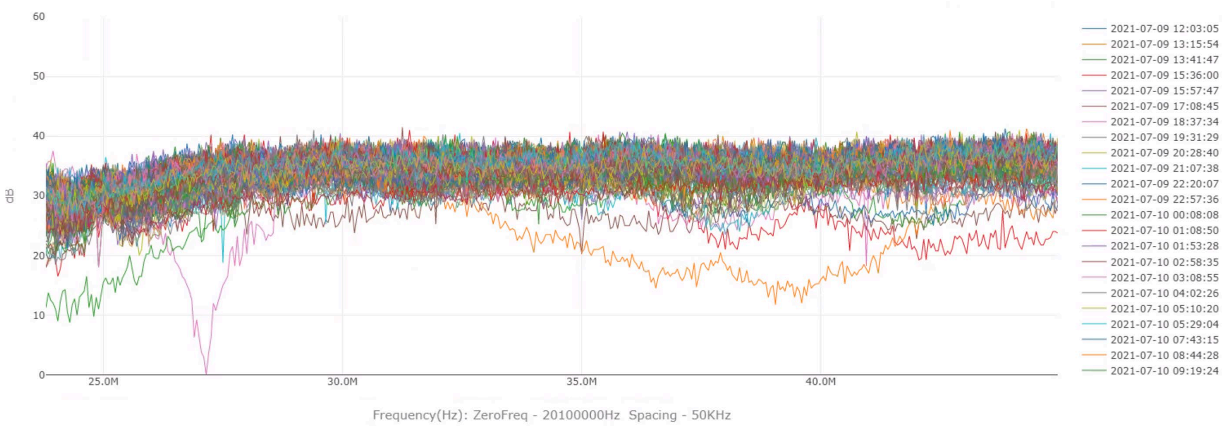
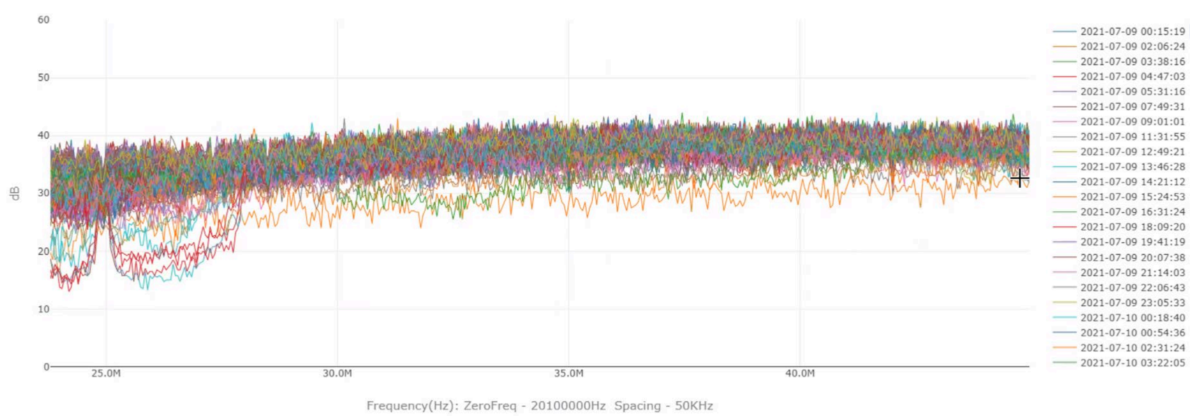


Figure 5 – Multiple US RxMER samples, OFDMA channel, Single CM (field)

The following RxMER captures show relatively tight value distribution but have many outliers throughout the capturing period.





Based on the observations of such variations, we develop statistical analysis methods to further extract information from the upstream RxMER data, which are discussed in the following section.

4. Statistical Analysis

When assuming that the RxMER values captured from each subcarrier are following normal distribution while the impairments are absent, it is helpful to use percentiles, variance, skewness, and kurtosis calculated from the RxMER values over time to summarize the distributions and provide insights into the behavior of the impairments. Each of the statistical calculations is discussed in the following sections. We also discuss initial use cases of the statistical metrics, observations of RxMER time series on different frequency ranges, and why the data suggests that PMA is necessary for the robustness of the upstream OFDMA deployment.

4.1. Percentiles

In order to analyze the upstream OFDMA RxMER values captured over time for each CM, we calculate different percentile values from each of the subcarrier's historical RxMER values. We calculate 1%, 2%, 5%, 10%, 15%, 20%, and 50% values from all the RxMER values captured on each subcarrier and include the minimum and maximum values to illustrate the range and distribution densities. By calculating the delta values between the percentile traces, one can automatically identify frequency ranges that may need attention as well as intermittent impairments on the data captured over time, as highlighted in the following figures.

For example, when we calculate the deltas between the minimum value, 1%, 5%, 15%, 50%, and the maximum value, and observe large differences between 1% and 5% traces while seeing smaller delta values between the other percentile traces, such as the condition indicated in the third highlight in the following figures, an intermittent impairment can be identified. On the other hand, if the percentile values are evenly distributed but the averaged percentile delta values are high, such as the condition shown in the first highlight, a persistent impairment can be identified.

Figure 6 – RxMER over time with persistent impairments and intermittent impairments

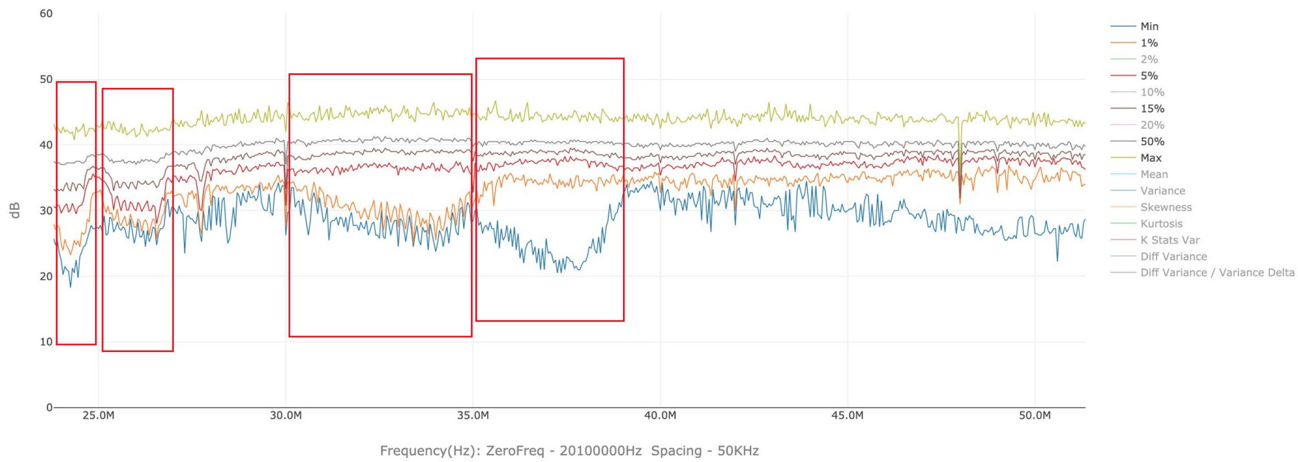


Figure 7 – Percentiles and min, max values calculated for each subcarrier

4.2. Variance over Time

It is well known that variance is a statistical measurement of the spread between numbers in a dataset. In the application we developed for this research, other than calculating the percentiles we calculate the variances of the RxMER values captured from each subcarrier and visualize the variance calculation results over frequency.

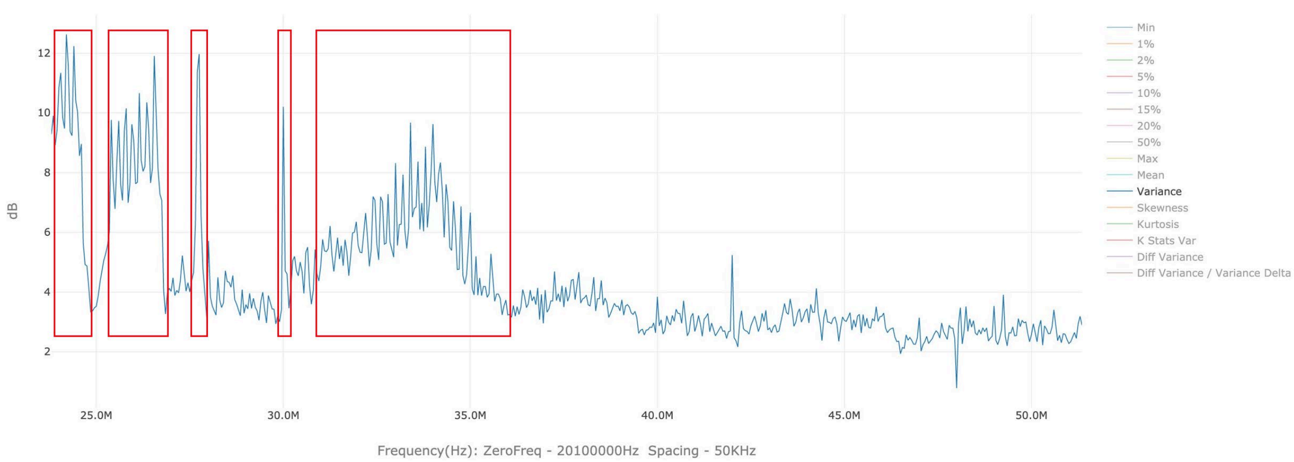


Figure 8 – Variance values calculated for each subcarrier

Relatively high variance can indicate that the RxMER values captured from the subcarriers are unstable over time. However, variance by itself would not provide sufficient information to differentiate persistent impairments and intermittent impairments/events, as both can cause instability of subcarriers' RxMER values which leads to high variance values. By defining thresholds for the variance, it can be used as the first step of selecting unstable subcarriers based on their upstream RxMER values captured over time. And in order to extract sufficient information for categorizing the impairments observed over time at high-level, such as identifying persistent impairments versus intermittent impairments, we introduce skewness and kurtosis calculations of the RxMER values from each subcarrier.

4.3. Skewness over Time

Skewness is a measure of the asymmetry of the probability distribution. The value of skewness can be positive, zero, or negative. Positive skewness indicates that there is more weight in the right tail of the distribution, whereas negative skewness indicates the opposite.

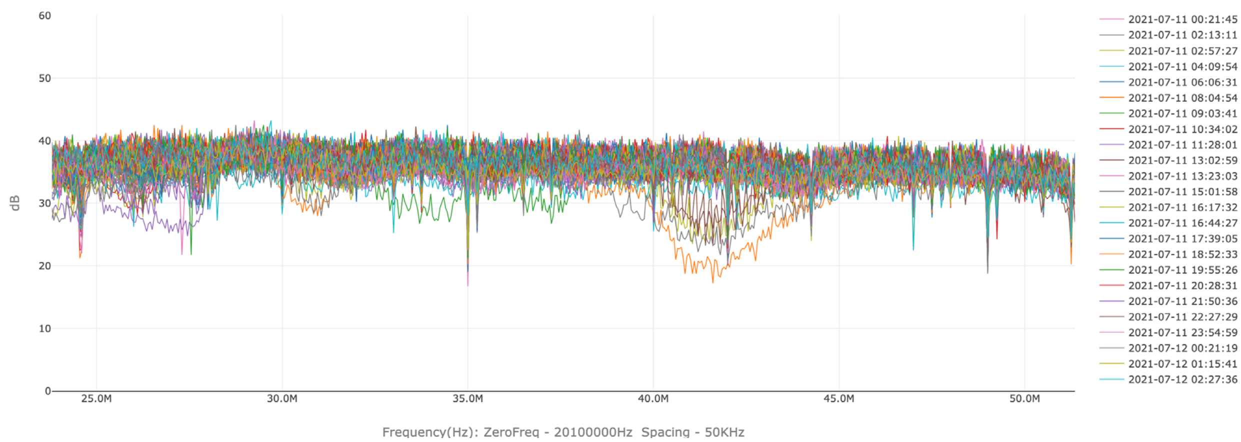


Figure 9 – A CM with persistent impairments and intermittent impairments

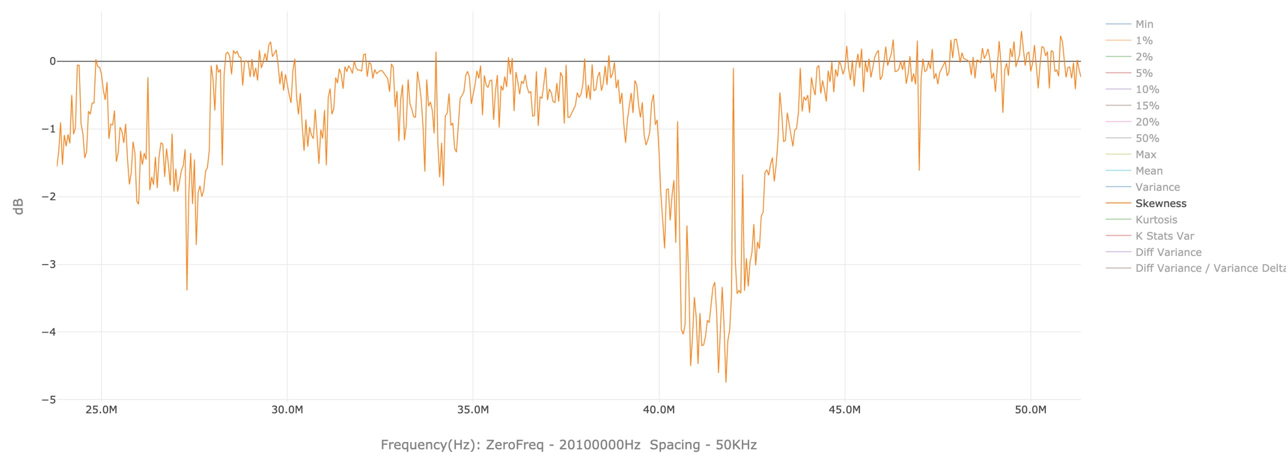


Figure 10 – Skewness calculated for each subcarrier

Considering that the distribution of a subcarrier's RxMER values over time should be approximately symmetrical when no impairment is present or only persistent impairments present, large absolute skewness values (especially when the skewness is negative) can be used to identify subcarriers affected by intermittent impairments/events.

Combining skewness with variance, it can be inferred that when both of the variance value and the absolute skewness value are large, the subcarrier is primarily being affected by intermittent impairments;

and when the variance value is large and the skewness value is close to zero, the subcarrier is primarily being affected by persistent impairments.

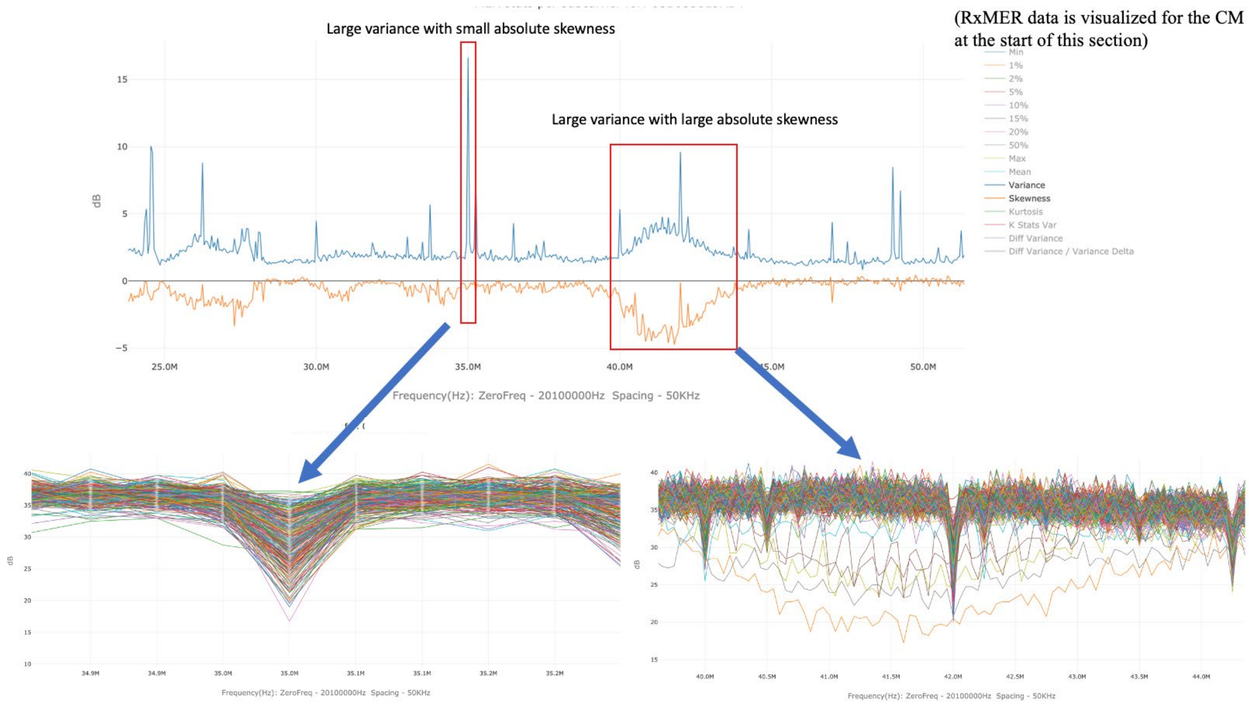


Figure 11 – Persistent and Intermittent Issues identified by variance & skewness

The bottom of Figure 11 shows that a persistent issue presents and is identified by using the combination of variance and skewness at 35 MHz of the spectrum. The variance of this subcarrier's RxMER values is large, however, the skewness is close to zero based on the calculation and the observation that the RxMER values are evenly distributed under the impairment.

Another impairment between 40 MHz and 44 MHz is identified as an intermittent issue since the variance values and absolute skewness values are large. The behavior of both of the identified impairments can be further confirmed when we visually check from the bottom of the 3-dimensional graph (Figure 12) generated from the upstream RxMER samples captured over time.

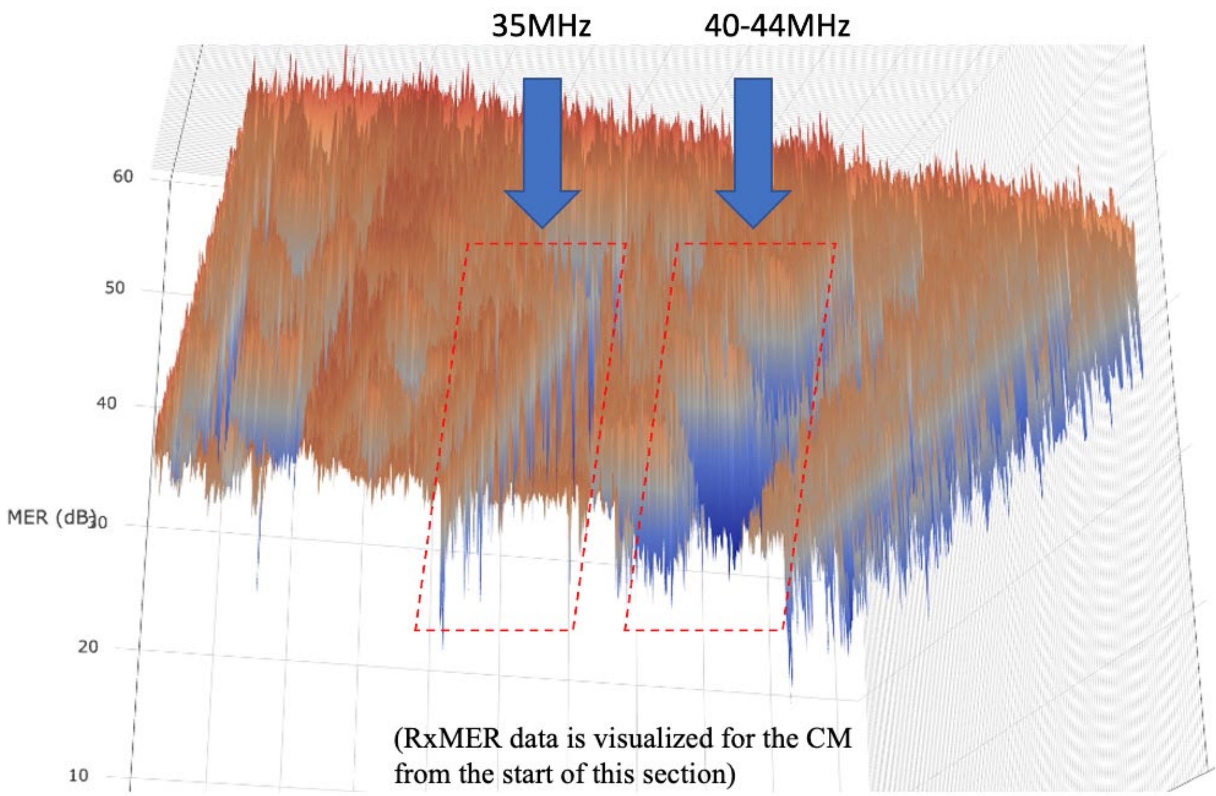


Figure 12 –3D graph of the RxMER samples captured over time

At 35 MHz, the impairment is persistent. It constantly affects the RxMER values and is almost captured by every RxMER sample. On the other hand, the impairment between 40 MHz and 44 MHz can be considered intermittent since it only presents occasionally in the 300 RxMER samples visualized in the 3D graph.

4.4. Kurtosis over Time

Kurtosis is a measure of the “tailedness” of the probability distribution. Higher kurtosis corresponds to greater extremity of deviations, which can be correlated with variance and skewness to further confirm if an identified impairment/event captured by the RxMER data is persistent or intermittent.
In our application, we calculate excess kurtosis (the value is 0 when the distribution is normal) of RxMER values over time for each subcarrier. From the observations, the combination of large absolute skewness values and large kurtosis values emphasizes the intermittent behavior of an impairment. This can be a promising technique to process a large number of RxMER samples captured from each CM over time, remove the noise, and filter out information that can be used for potential automatic anomaly detection in OFDMA, as demonstrated in Figure 13.

(RxMER data is visualized for the CM at the start of this section)

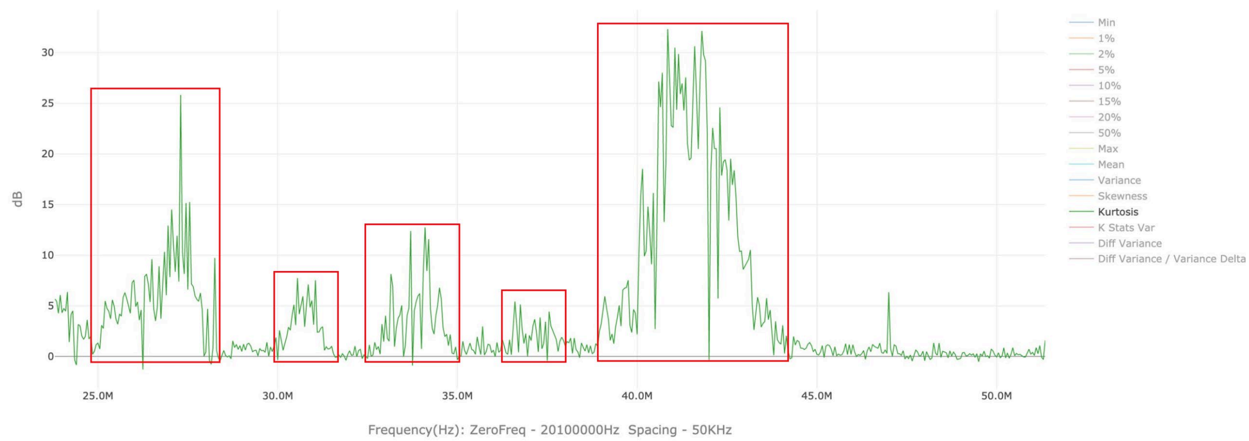


Figure 13 – Intermittent issues identified by kurtosis

4.5. RxMER Time Series on Different Frequency Ranges

Another way to analyze RxMER data is to view the variations of small frequency range (practically a multiple of the subcarrier size). Here we choose a unit of 1 MHz, which for the 50 kHz spacing is 20 subcarriers. The idea is to take an average of the RxMER for those twenty subcarriers and then plot that average value as it changes over time. We do this for every 1 MHz of the spectrum, and then certain patterns become more apparent. In order to observe RxMER value changes over time sequentially, we calculate the average RxMER values of each 1 MHz frequency within the channel and visualize the averaged RxMER values with the sample capture timestamps (as shown in Figure 14).This visualization method provides insights into the unstable nature of the upstream and helps us identify issues that can be potentially identified by algorithms we may develop.

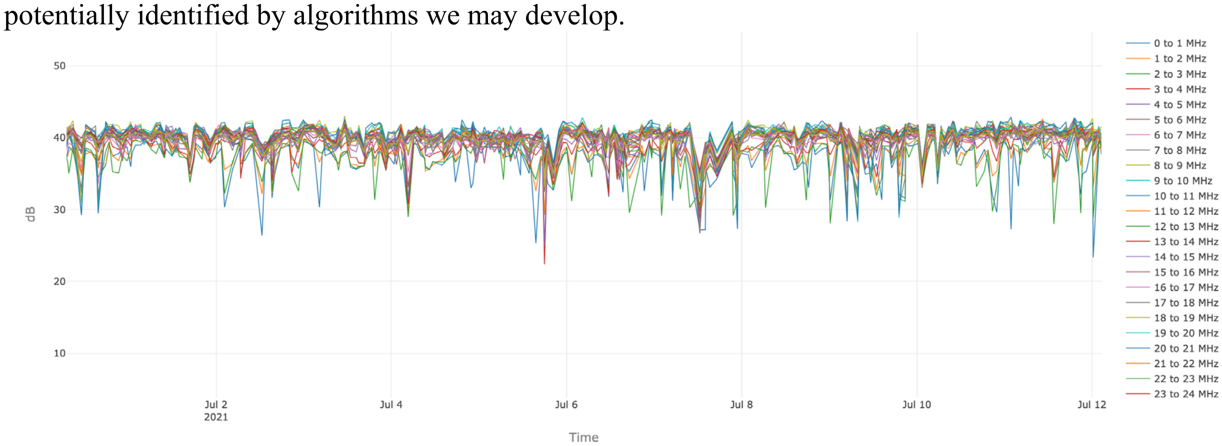


Figure 14 – Averaged RxMER time series for one CM

One of the patterns (see Figure 15) we observed is periodical changes across the whole channel over the course of the day, and these varying RxMER patterns repeat every day. The variation is significant, anywhere from 5 to 10 dB.

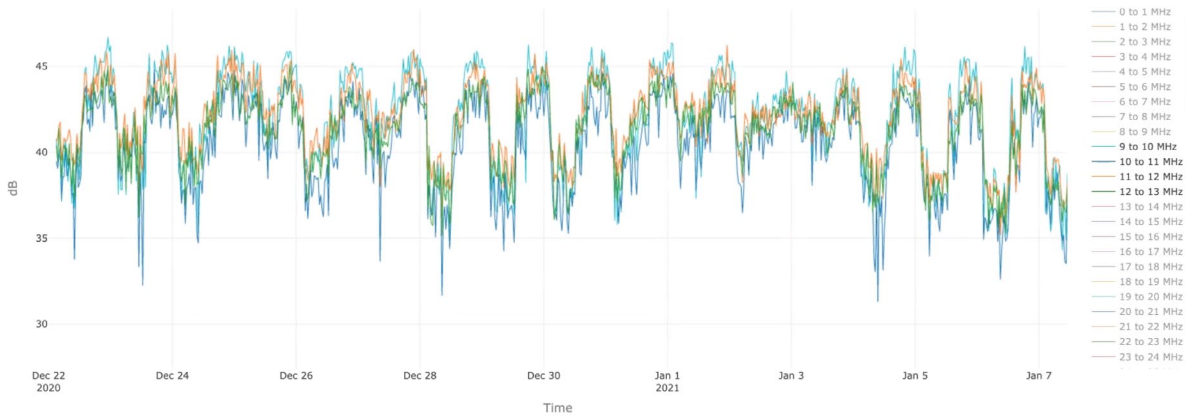


Figure 15 – Recurring daily variation in RxMER

In another example, (see Figure 16), an intermittent issue happened on July 5th can be identified between 11 MHz and 15 MHz of the OFDM channel (34.9 MHz – 38.9 MHz), where there is a sudden drop in the average RxMER values by about 18 dB.

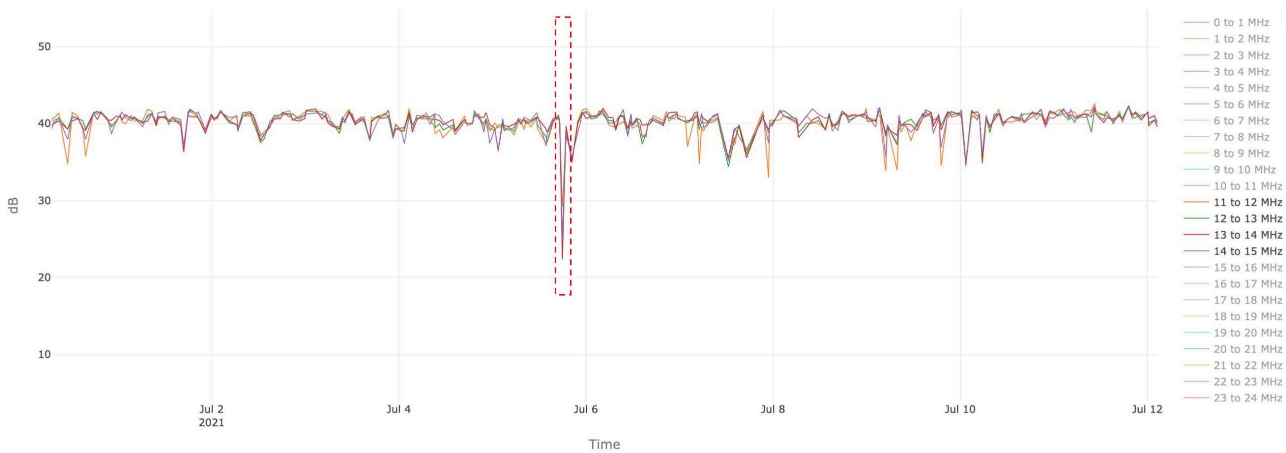


Figure 16 – Intermittent issue observed on RxMER time series

We can then target the problematic RxMER sample using the timestamp, as shown in the top half of Figure 17. Comparing the RxMER sample that has issues (the yellow trace) with a clean sample captured on the same day (the red trace), it indicates that an intermittent issue happened and significantly affected the result of the RxMER measurement.

In addition, because of noise funneling in the upstream, such an issue may also be captured by the measurement of upstream RxMER on a different CM around the same time, as shown in the bottom half of Figure 17.

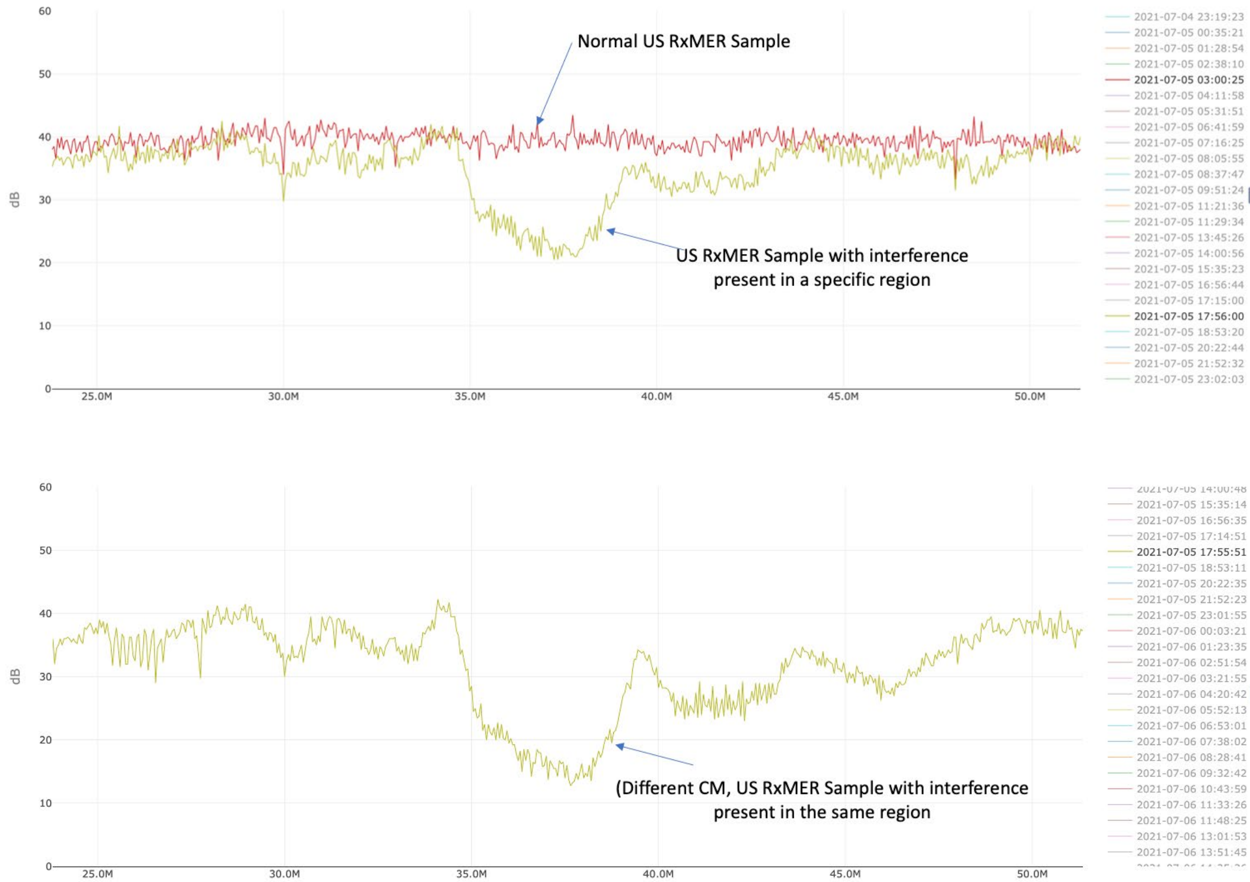


Figure 17 – A problematic frequency region on multiple CMs

To provide another example of such intermittent issues being highlighted by RxMER time series, as shown in the top of Figure 18, a significant event is identified using the data collected from a different CM where all subcarriers had a 30 dB drop in their RxMER values. This outlier sample is as shown in the bottom half of Figure 17. The cause of this outlier sample is under investigation, this could be a CMTS measurement bug or it could be the nature of the plant noise manifesting itself as a low RxMER.

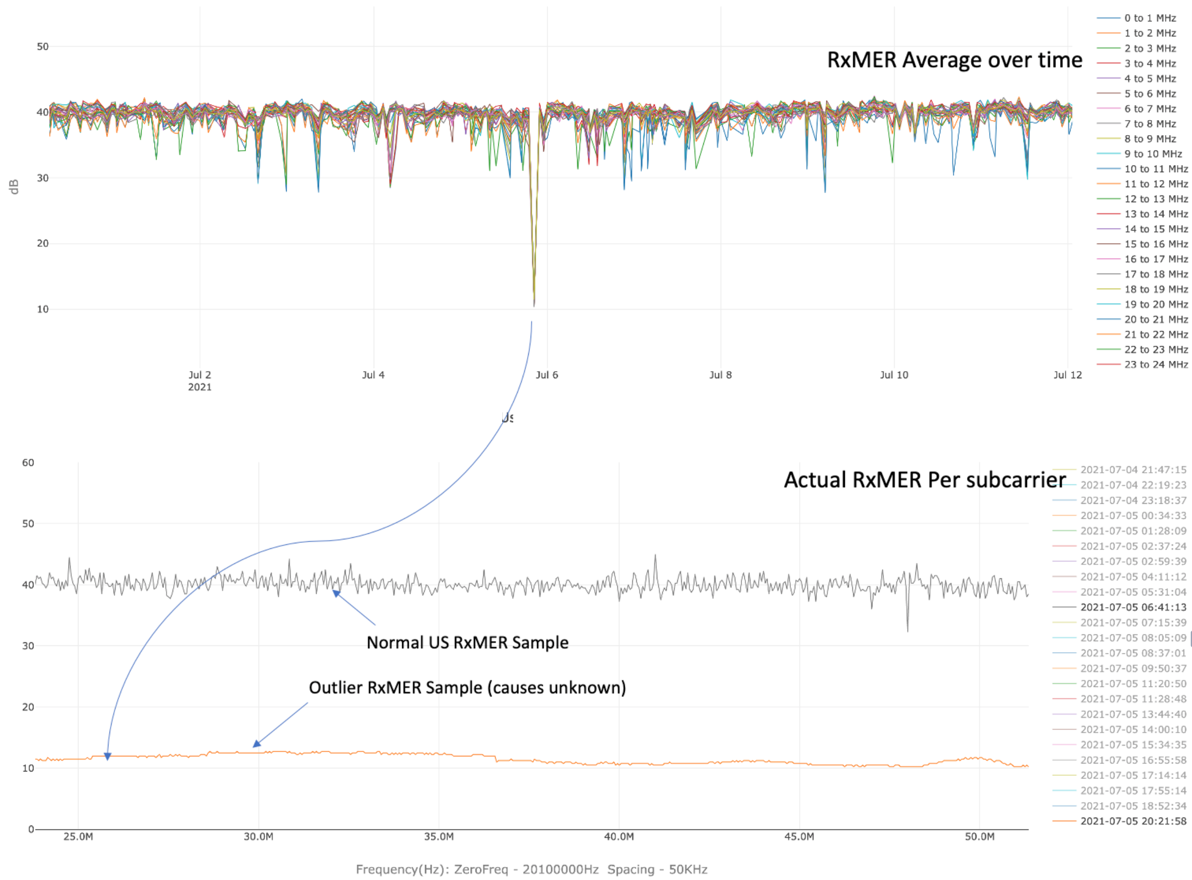


Figure 18 – A different intermittent issue observed on RxMER time series

In another example, one of the observations we made were sudden drops in average RxMER across the whole channel. For example, in Figure 19, we can see on Dec 27th the RxMER for the CM drops down by 10 dB, and then a few days later goes higher and fluctuates between two or three levels.

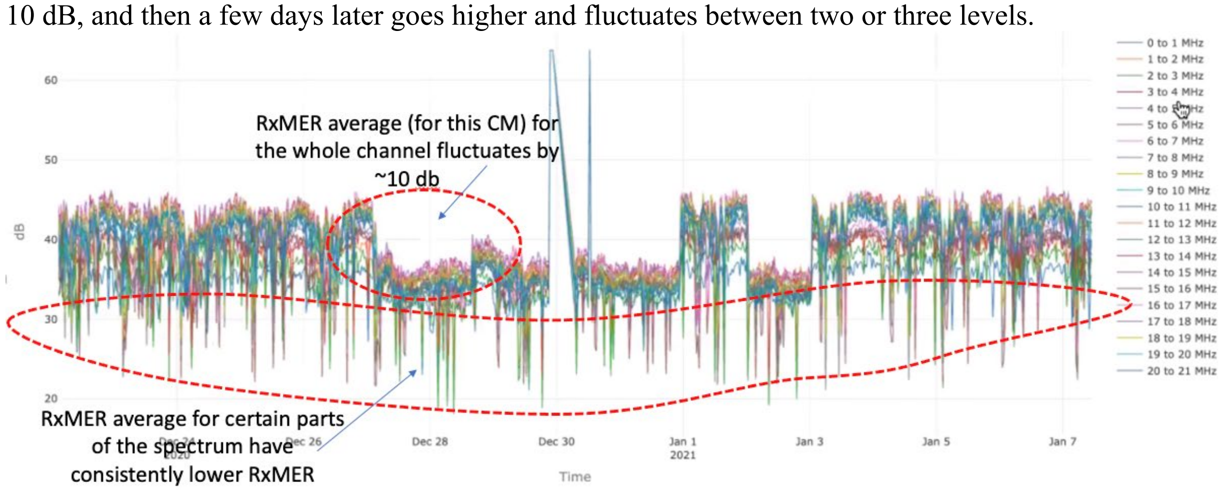


Figure 19 – RxMER Variation across time.

By using the average RxMER time series data, intermittent issues (both unknown issues and repetitive issues) can be identified with low computational cost. For the goals of proactive network maintenance (PNM), such identified issues can be characterized and categorized for the operators to be aware of and potentially reduce maintenance time and cost. In addition, when the root causes of the issues are located and isolated, the features of these issues can be fed into advanced models such as machine learning based classifiers for automatic anomaly detection and root cause inference.

4.6. Clustering Analysis on Statistical Measures

In order to analyze the features extracted from the upstream RxMER data collected over time using variance, absolute skewness, and kurtosis, and research how the impairments can be automatically isolated and categorized, we first calculate the statistics for all of the subcarriers of CMs on multiple OFDMA channels, and then generate a 3-dimensional graph to show if the statistics of the subcarriers' RxMER data are naturally clustered in 3D space.

As shown in Figure 20, there are two main clusters in the 3D graph, one of which has relatively large variance values and small absolute skewness and kurtosis values, whereas another cluster has small values in variance but relatively large values in both absolute skewness and kurtosis.

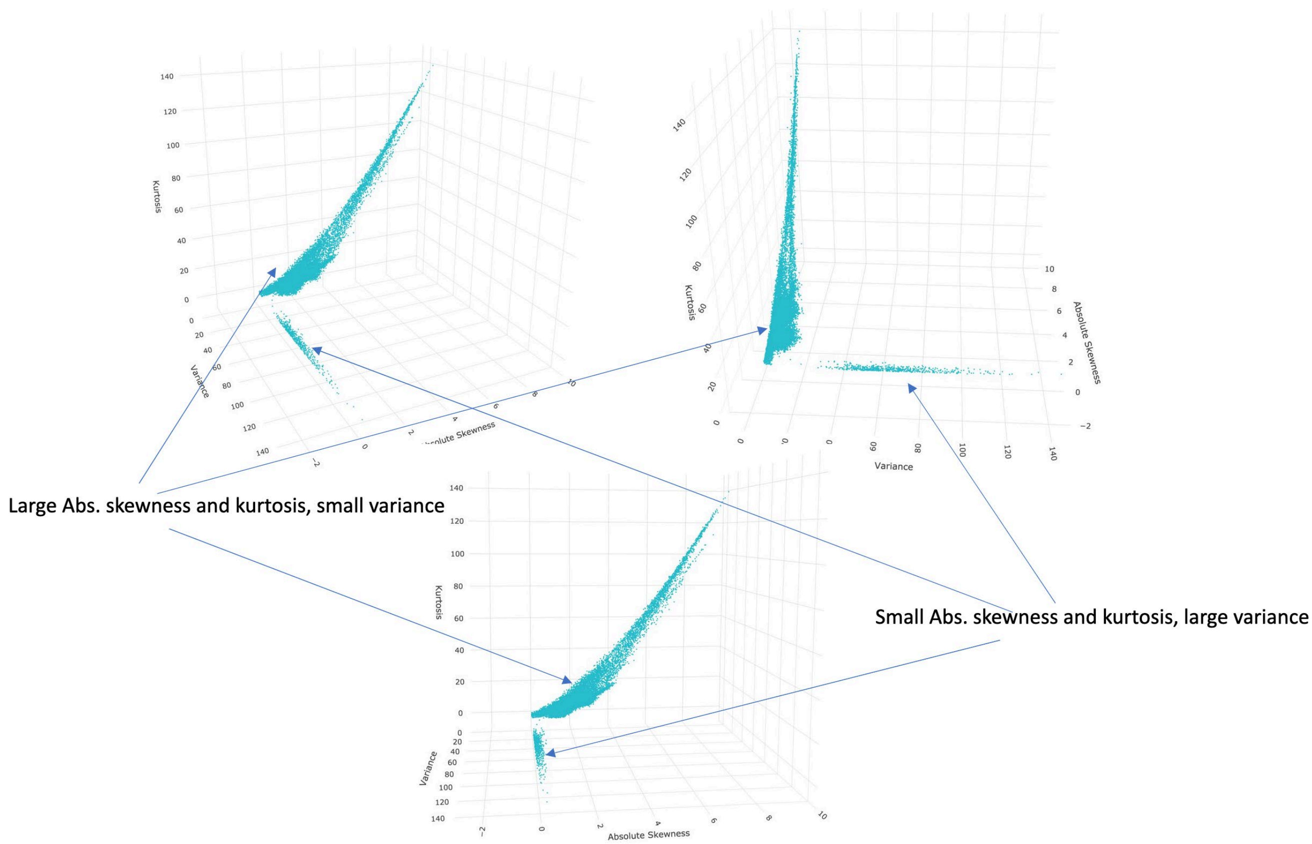


Figure 20 – 3D graph of variance, absolute skewness and kurtosis values

It can also be inferred that the cluster that shows significance in absolute skewness and kurtosis suggests positive relationship between the two statistics. In other words, large kurtosis and absolute skewness both can be used as criteria to identify intermittent impairments on the OFDMA channel, and they potentially reinforce each other.

In addition, this analysis helps identify thresholds that can be applied to the statistics in creating a threshold-based method to automatically detect subcarriers under impairment. Also, centroids can be calculated for these clusters to categorize the impairment types.

More research is ongoing around understanding the impairment types in upstream OFDMA channels. Calculating the statistics of upstream RxMER values over time can be a promising feature extraction method for machine learning models. For example, a 1-dimensional convolutional neural network can use the normalized variance, skewness, and kurtosis traces over frequency as input taken from 3 different channels. A regression model (see [ANOMALY DETECTOR ICPHM 2020]) can be developed to automatically localize and classify impairments observed by RxMER data in upstream OFDMA.

4.7. Time Clustering – PMA Use Case

By analyzing the upstream RxMER data collected from the field, an important note can be made that the condition of the OFDMA channel can vary significantly from moment to moment, which can lead to frequent IUC shifting if the CMTS is enabled to automatically adjust upstream IUCs for CMs based on FEC error rates and RxMER measurement results. To automatically create a set of IUCs that can provide optimal operational robustness and channel capacity, upstream PMA can be adopted. The upstream PMA algorithm uses time clustering to group all upstream RxMER samples collected from all CMs over time in order to automatically design optimal IUCs, which has been proven to be beneficial in field trials.

There are 2 algorithms that are implemented in upstream PMA which we call algorithm 2b and 2c, see [US PMA SCTE 2020]. Algorithm 2b is more aggressive in capacity improvements compared to algorithm 2c, whereas algorithm 2c designs more robust IUCs. To provide examples, we calculate and visualize the time each CM spends on each IUC to provide insights on how the CMs may utilize the IUCs based on simulated CMTS criteria. As shown in Figure 21, IUCs created by algorithm 2b are used by the CMs evenly in time, as the IUCs are designed to prioritize channel capacity gains. In Figure 22, CMs use IUC 5 and IUC 6 primarily, as the bit loadings of the IUCs are lower to ensure robustness and fewer IUC shifts.

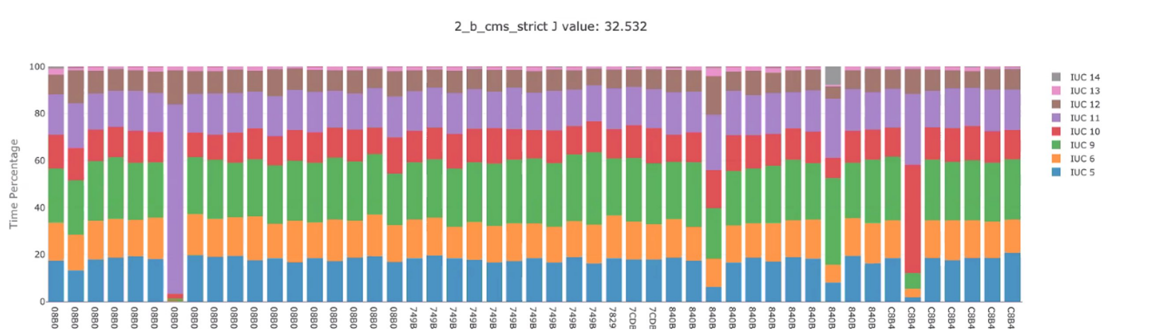


Figure 21 – IUC usage time distribution per CM (method 2B)

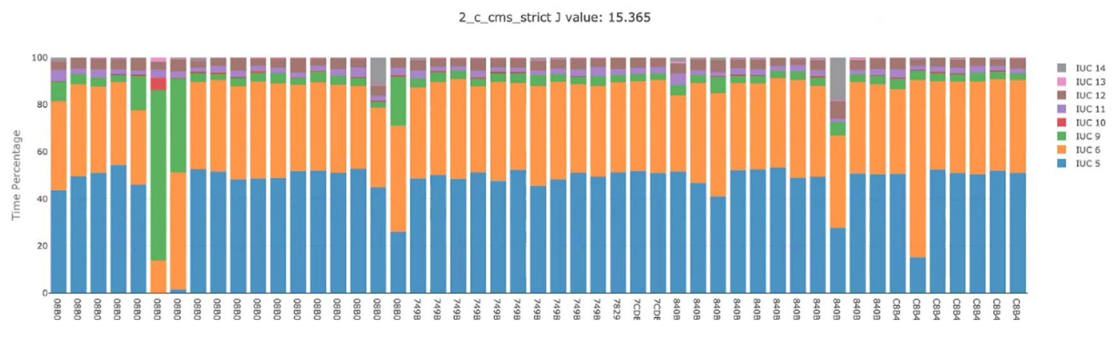


Figure 22 – IUC usage time distribution per CM (method 2C)

5. Data Analysis

We have observed some interesting patterns in the OFDMA channel data collected by the CMTS on modems in the field. These include "outage" patterns on a particular modem on a particular OFDMA channel, or invalid RxMER values for a CM across the channel etc. To categorize the issues in data reporting, we label them as "measurement discontinuities".

5.1. US RxMER Measurement Discontinuities

The first pattern that we see is that some modems have no RxMER entries for large parts of the data collection timeframe. As an example, in the figure below, during a data collection period of two weeks, we see that for a period of four days the modem does not respond and so the CMTS has no RxMER samples for that modem. This may be caused by but true outage or an issue at the cable modem or it could also be that the user just turned the cable modem off. Some data analysis would need to be done to correlate the outage with say data traffic on the modem and tease out the root cause of this outage. A second issue that we observe is when the CMTS reports value of 0xFF for all the subcarriers off that modem. Based on some discussions with the operator on the status we believe that this condition is when the CM is in partial service on that OFDMA channel but is maintaining data connectivity with the CMTS on the other channels.

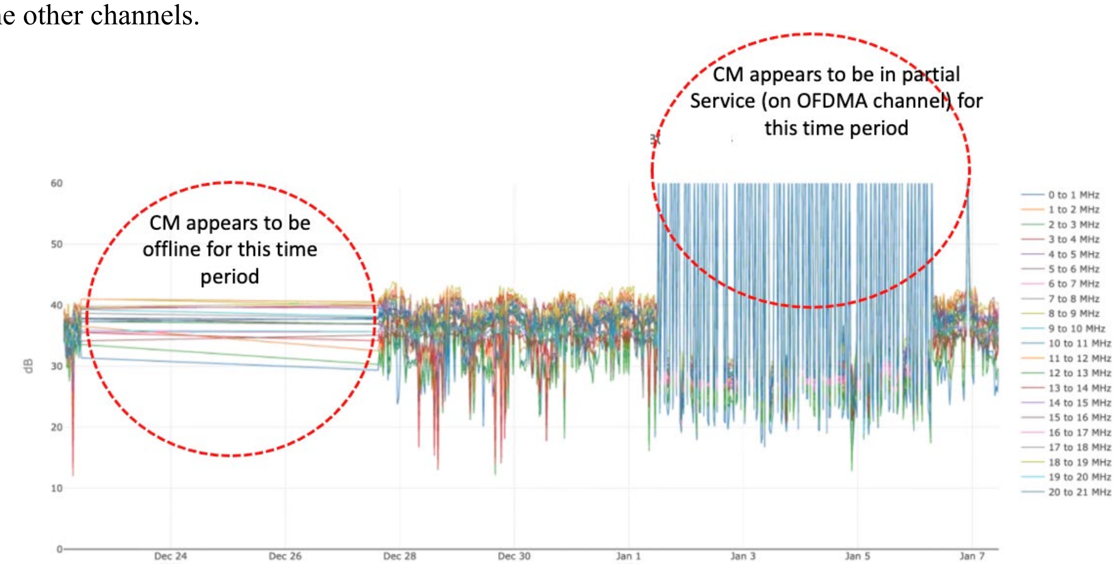


Figure 23 – One CM offline or in partial service

5.1.1. CM Outage Issues

Based on examples we saw in the previous section we know that multiple modems have outages at different points in time. This leads us to the next step of analysis by aggregating all such outage data from modems on the same node. The idea is to observe each of the individual CM outages on the node and try to correlate them over time to identify any meaningful events on the network.

Here we aggregated all the offline outage events (i.e., no RxMER samples) from the modems and are plotting them as a bar graph over time. Given that the data collection frequency was every hour we consider each gap in the data set as one event for that one modem. We calculate the number of such outage events and plot them on the graph below as an aggregate of all modems in the service group. The graph below shows that between Dec 24th and 28th there were many hour-intervals where up to 13 to15 modems in the service group were unable to respond to the RxMER request and work offline during that interval

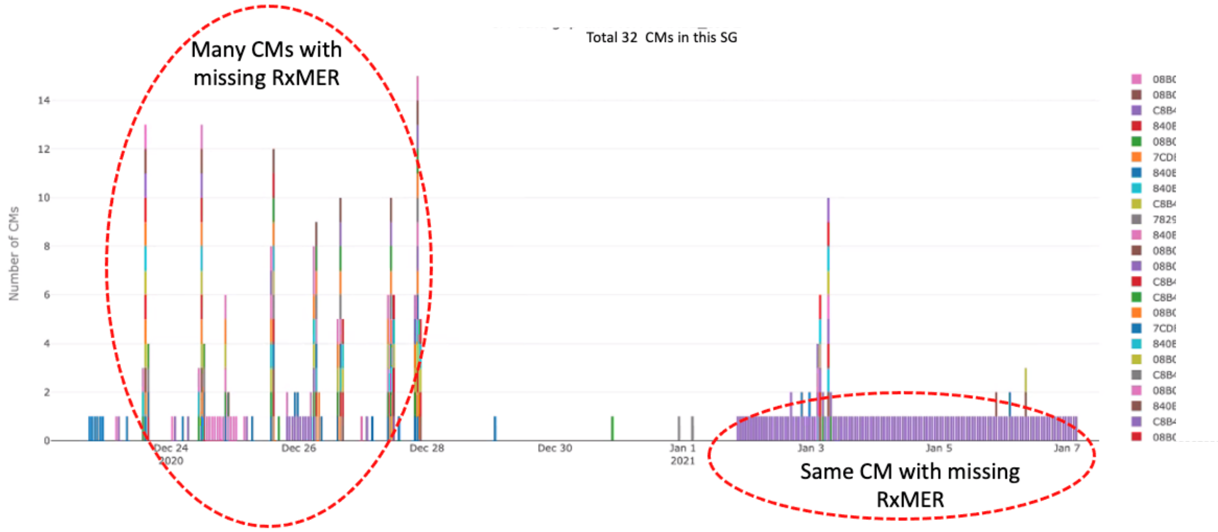


Figure 24 – Num of CMs Missing Data Samples across SG

We also looked at the amount of time each modem was missing data and plotted a histogram of the outages. The graph below shows the number of events which lasted from ~50 minutes to ~3 hours.

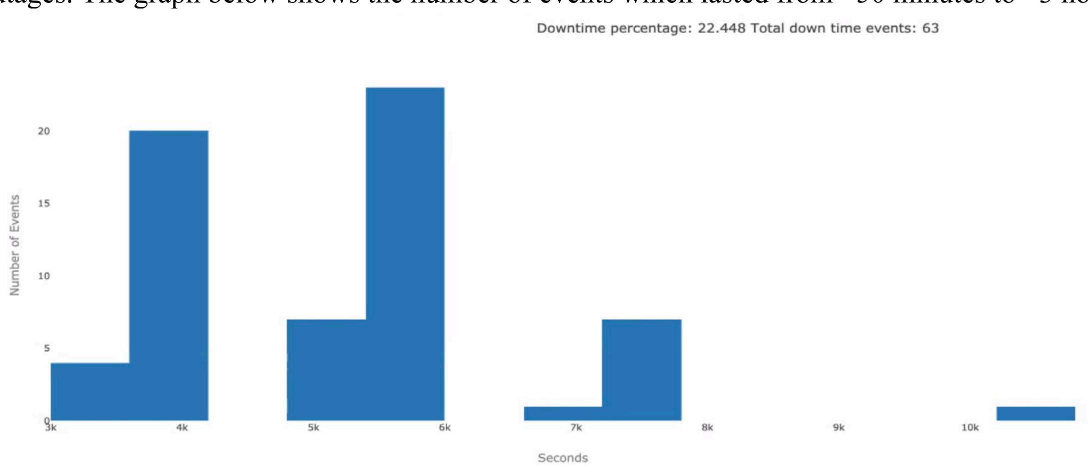


Figure 25 – Characterizing outage times

5.1.2. CM Invalid RxMER

The second type of measurement discontinuity are the events when the CMTS returns all 0xFF values (i.e., invalid samples) for US RxMER from the modems. Any collection interval which has a US RxMER file but with the data set to all 0xFFs counts as one event for that one modem. We calculate the number of such outage events and plot them on the graph below as an aggregate of all modems in the service group. The graph below shows that between Jan 1st and Jan 7th, there were many hour-intervals a few different modems (different colors for the actual CM MAC address), up to 2 at a time, were unable to respond to the CMTS request appropriately to get valid RxMER measurements. In particular two CMs, in green and gray looked to be struggling during that week.

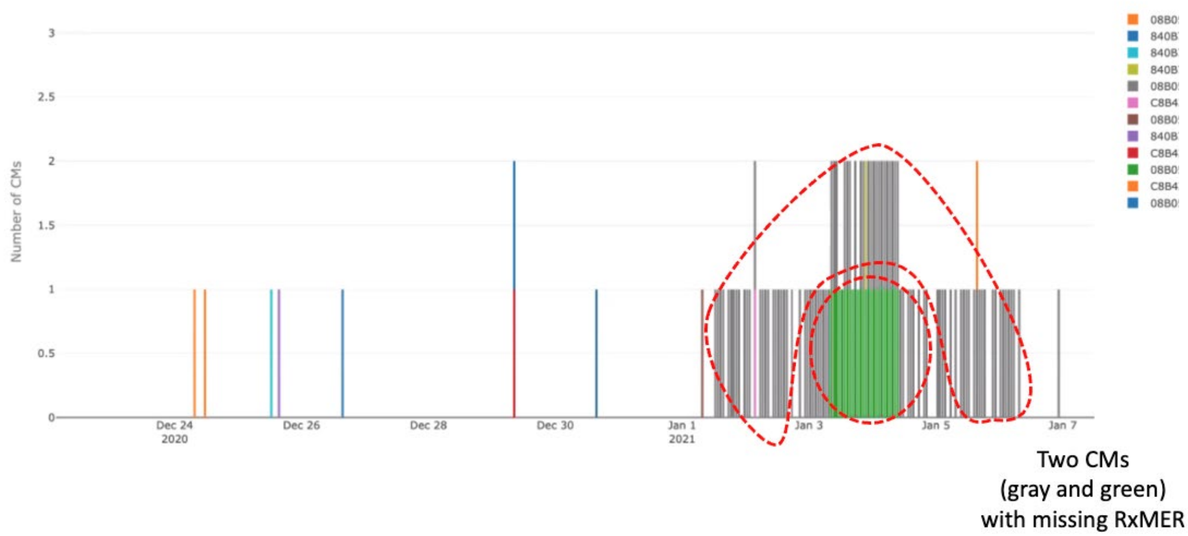


Figure 26 – Num of CMs offline across the whole SG (FF)

5.2. Health Score

Operators have identified a need to develop a set of metrics from the network which reflect how they affect the customer's network connection (speed, reliability, latency). Operators want to identify which factors have an impact on the service. Any network health metric needs to be ultimately linked to the customer impact.

A health score allows an operator to discern which alarms/events from the network are relevant. An operator would like to use a health score to prioritize problems and help triage problems and identify which areas/problems need truck-rolls.

One can think of various components to the health score, including downstream metrics upstream metrics node level metrics etc. Here we are just focused on the upstream component of an overall health score. We look at the individual CM level score and add a service group or node level score. At this point these upstream scores themselves could consist of multiple components though for this paper we are focused only on the RxMER data.

5.2.1. CM Score Upstream

Each of the parameters that we discussed above such as variance skewness, kurtosis, percentile, could all have different thresholds or areas of interest where we could flag the particular modem. One can also look at FEC statistics or profile/IUC statistics, and combined them into a CM score

5.2.2. Node Score Upstream

One can also aggregate all the measured metrics from all the modems on a given node and come up with an aggregate node score. The benefit of this view is for an operator to prioritize how a particular node is doing and how healthy is a node as compared to other nodes on the plant. This gives and operator a prioritized list of nodes which they can then work through the top issues.

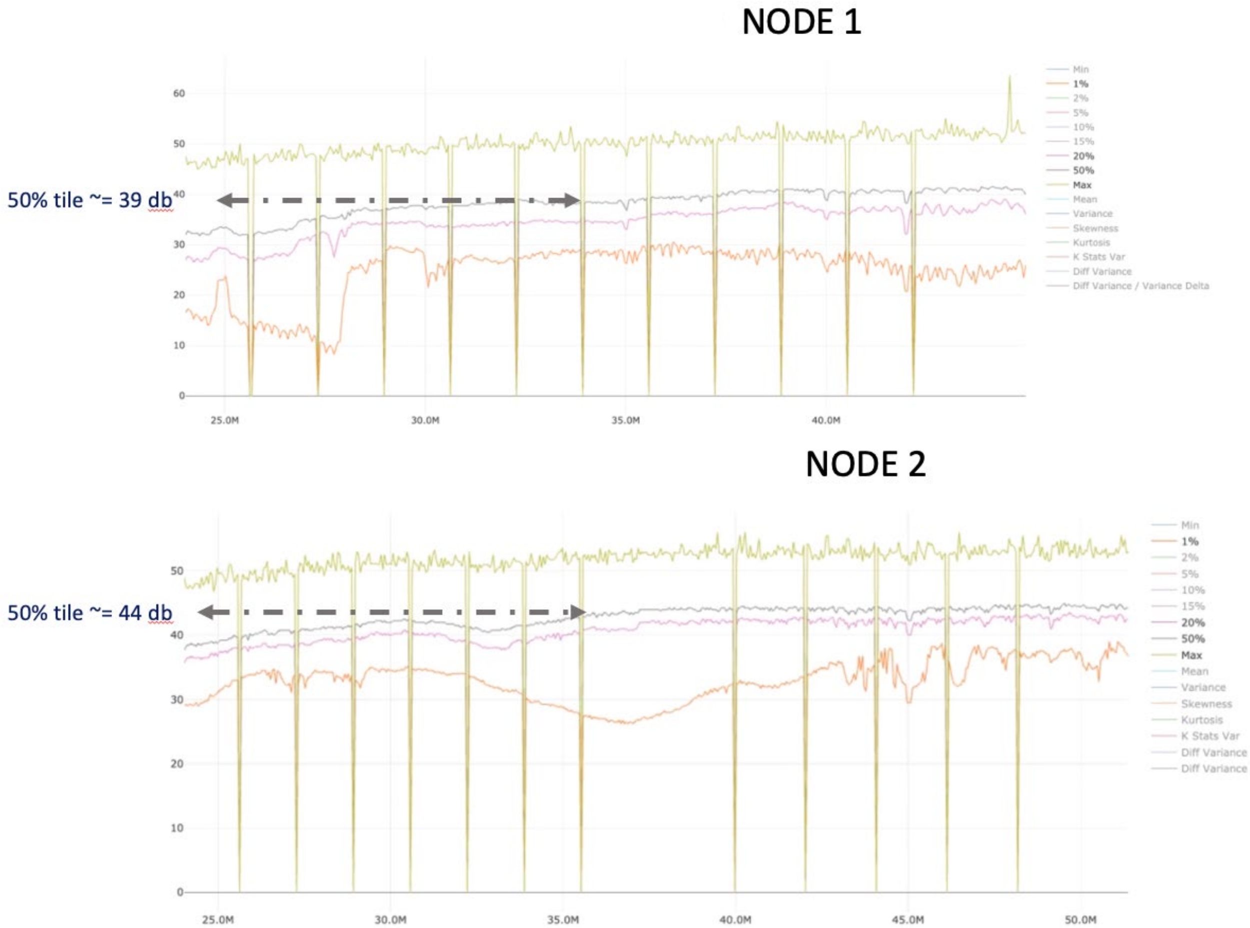


Figure 27 – All CMs All RxMER samples, Percentile, Node 1 vs Node 2

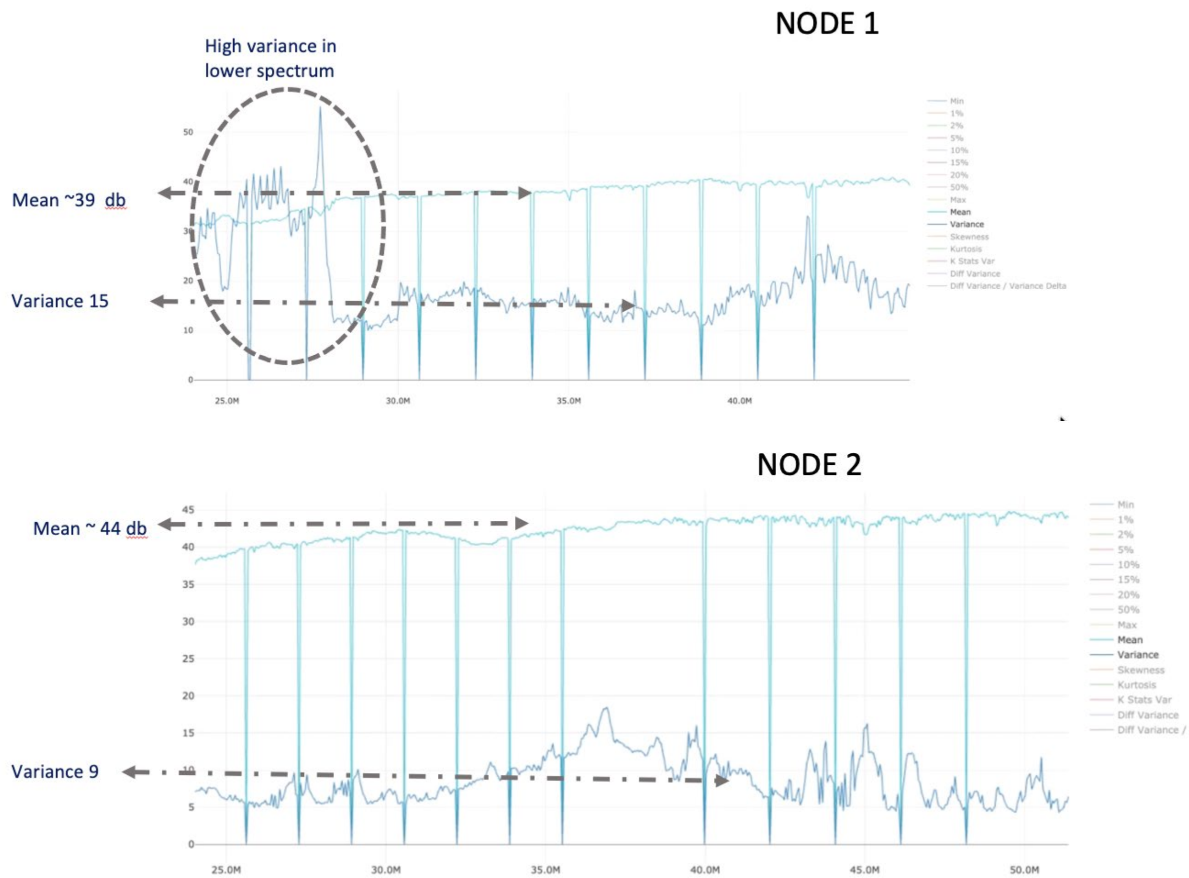


Figure 28 – All CMs All RxMER samples, Variance, Node 1 vs Node 2

5.2.3. US RxMER Data Analytics Application

CableLabs has developed an upstream RxMER data analysis application which can be used for a variety of functions [C3 CableLabs]. As shown in this paper one can visualize the RxMER of a cable modem, all the RxMER samples of a CM, all the RxMER samples of a node. There is also a table view off a service group where you can CB statistics off each individual cable modem. An operator can obtain the average variance, skewness, kurtosis in each part (frequency ranges) of the channel. An operator can also sort the CMs based on those values.

This is a very powerful tool in identifying the "outlier" CMs in the node/segment and then use that information for further analysis/ triaging of truck rolls etc.

CM MAC	Spacing	First Active Sub-carrier	First Active Sub-carrier Frequency	Interface Number	SG Name	Average Variance (lower spectrum)	Average Variance (higher spectrum)	Average Skewness (lower spectrum)	Average Skewness (higher spectrum)	Average Kurtosis (lower spectrum)	Average Kurtosis (higher spectrum)	Downtime Percentage	Gap Time Percentage	Show Data
08B0	50	74	24025000	34259153	AM	182.14	161.22	0.92	0.93	-0.38	-0.3	22.448	33.036	Show Data
08B0	50	74	24025000	34259153	AM	46.21	32.62	1.81	1.79	6.68	7.21	6.633	4.719	Show Data
7CDB	50	74	24025000	51003593	AM	29.22	25.73	1.54	1.43	9.03	6.4	4.081	2.679	Show Data
08B05	50	74	24025000	34259153	AM	27.42	24.05	-1.56	-2.21	5.34	8.64	0	0	Show Data
7CDB	50	74	24025000	51003593	AM	26.93	35.09	-0.45	-0.44	-0.54	-1.08	0	0	Show Data
94917	50	74	24025000	51003593	AM	26.02	24.24	0.38	0.29	6.28	3.55	1.148	0.256	Show Data
840B7	50	74	24025000	34259153	AM	25.03	18.81	-0.33	-1.03	2.16	5.93	0.255	4.591	Show Data
7CDB	50	74	24025000	51003593	AM	23.55	29.39	-0.45	-0.37	-0.25	-1.12	0	0	Show Data
840B7	50	74	24025000	51003593	AM	23.34	22.06	0.07	0.17	4.94	2.01	0.765	0.255	Show Data
840B7	50	74	24025000	34259153	AM	22.81	25.35	-1.17	-2.3	3.71	9.84	0	2.041	Show Data

Showing 1 to 10 of 113 entries

Previous 1 2 3 4 5 ... 12 Next

CM MAC	Spacing	First Active Sub-carrier	First Active Sub-carrier Frequency	Interface Number	SG Name	Average Variance (lower spectrum)	Average Variance (higher spectrum)	Average Skewness (lower spectrum)	Average Skewness (higher spectrum)	Average Kurtosis (lower spectrum)	Average Kurtosis (higher spectrum)	Downtime Percentage	Gap Time Percentage	Show Data
08B0	50	74	24025000	84570313	AM	4.15	2.37	-0.64	0.8	4.39	14.56	0	0	Show Data
7CDB	50	74	24025000	84570313	AM	7.03	4.42	-0.88	-0.65	2.7	3.16	0	0	Show Data
08B05	50	74	24025000	84570313	AM	7.04	4.84	-0.82	-0.87	2.67	4.17	0	0	Show Data
1CB0	50	74	24025000	84570313	AM	7.07	4.64	-0.85	-0.85	2.7	4.13	0	0	Show Data
840B7	50	74	24025000	84570313	AM	7.36	5.25	-0.92	-1.08	2.92	4.88	0	0	Show Data
08B05	50	74	24025000	84553937	AM	7.53	3.2	-0.39	0.8	0.54	2.92	0	0	Show Data
C8B42	50	74	24025000	84570313	AM	7.62	5.57	-0.97	-1.21	3.1	5.47	0	4.592	Show Data
94917	50	74	24025000	84570313	AM	7.74	5.34	-1	-1.19	3.18	5.32	0	0	Show Data
94917	50	74	24025000	84570313	AM	7.75	5.54	-1	-1.25	3.18	5.69	0	0	Show Data
94917	50	74	24025000	84570313	AM	7.92	5.72	-0.99	-1.34	3.04	5.88	0	0	Show Data

Figure 29 – All CMs All RxMER samples, Variance, Node 1 versus Node 2

5.3. Future Work

There are future topics we will continue to research on in order to develop advanced methods for automatic upstream OFDMA anomaly detection for cable operators to gain more visibility into the upstream OFDMA performance, potentially guide truck rolls and reduce network maintenance cost. These topics include defining anomaly categories in the upstream, anomaly detection methods, and correlations of different measurement data.

5.3.1. Defining Anomaly Categories in the Upstream

We observe many common impairments in the upstream RxMER such as the ingress-like noise observed between 15 and 27 MHz of the OFDMA channel as shown in Figure 30. However, the sources of the impairments mostly remain unknown today. Apart from this, the types and features of the upstream impairments vary from interface to interface, node to node, and area to area. In order to categorize the observed anomalies in the upstream, one may define the labels of the impairments in a hierarchical way. For example, the two parent categories could be persistent impairment and intermittent impairment. Under each of the parent categories, there could be sub-categories such as ingress, echo, or loose connector etc.

Defining impairment categories in the upstream OFDMA can provide critical insight into understanding performance affecting issues and how they need to be prioritized. It also guides the development of feature extraction methods and anomaly detection methods.

Figure 30 – Impairments at lower frequencies

5.3.2. Anomaly Detection Methods

Several different types of anomaly detection methods can be applied to the upstream RxMER dataset. Threshold based algorithms can be used to label impaired subcarriers and differentiate intermittent issues and persistent issues; 1-dimensional convolutional neural networks can be used to detect patterns over time (RxMER time series) or over frequency; 1-dimensional convolutional neural networks that has multiple input channels can be used to take aggregated data as input, for example, variance, skewness, and kurtosis of each subcarrier's RxMER values over time to perform pattern recognition of the anomalies; 2-dimensional convolutional neural networks can be used to directly consume RxMER sample captures over time (the 2 dimensions are frequency and time) and identify events with their frequencies and timestamps, similar to recognizing objects on pictures; Recurrent neural networks can be used to recognize patterns on RxMER time series. All of these candidate methods depend on well-defined categories/patterns and labeled datasets.

5.3.3. Correlation of Different Measurement Data

While the upstream OFDMA RxMER data already provides much information for performance monitoring, correlating it with other measurement data could have enhanced benefits. Upstream Pre-Equalization data can provide insights into how the CMTS configures the CMs to compensate impairments or other channel attributes; upstream triggered spectrum captures can provide detailed views of the upstream transmission status; upstream FEC error rates can be correlated with RxMER captures over time to analyze how much a certain impairment/event is affecting the robustness of the communication. All of these can be interesting research topics in upstream data analytics.

6. Conclusion

As OFDMA becomes more and more widely deployed, the upstream RxMER measurement capability is becoming available and mature on the CMTSs and MSOs have started to leverage this capability to gain visibility into the status and performance of OFDMA. By visualizing the upstream RxMER captures and producing different views of data, information such as upstream impairments and RxMER variations can be captured to support and drive research of upstream data analytics. In this paper, we discussed how statistical methods can be used to analyze RxMER values captured over time from each subcarrier and extract useful information from noisy upstream datasets. Intermittent issues and persistent issues can be isolated and differentiated by calculating variance, skewness and kurtosis on the upstream RxMER data captured over time. And by plotting the statistics in 3-dimensional space, we discover clusters that can potentially be leveraged to support automatic impairment detection/categorization. Analyzing the RxMER data captured over time by creating average RxMER plots for different frequency regions can provide views from another aspect of the data. Combining the RxMER over time data with identified CM reporting issues, MSOs can gain more visibility into the history of OFDMA channels' performance on individual CM basis as well as at node level. As of the goals of future research of upstream OFDMA data analytics, it is possible that more methods for feature extraction can be developed to simplify and aggregate measurement information. And intelligent models can be built to localize and classify upstream OFDMA impairments accurately and efficiently.

Abbreviations

3D	three-dimensional
bps	bits per second
CM	cable modem
CMTS	cable modem termination system
dB	decibel
DOCSIS	Data-Over-Cable Service Interface Specifications
FEC	forward error correction
HFC	hybrid fiber-coax
Hz	hertz
IUC	interval usage code
ISBE	International Society of Broadband Experts
kHz	kilohertz
MHz	megahertz
MSO	multiple-system operator
OFDM	orthogonal frequency division
OFDMA	orthogonal frequency division multiple access
PMA	profile management application
P-MAP	probe map
PNM	proactive network maintenance
RxMER	receive modulation error ratio
SCTE	Society of Cable Telecommunications Engineers
SC-QAM	single channel quadrature amplitude modulation
TFTP	trivial file transfer protocol

Bibliography & References

[DOCSIS PHYv3.1] DOCSIS 3.1 Physical Layer Specification, CM-SP-PHYv3.1-I18-210125, January 25, 2021, Cable Television Laboratories, Inc.

[DOCSIS MULPIv3.1] DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv3.1-I21-201020, October 20, 2020, Cable Television Laboratories, Inc.

[US PMA SCTE 2020] K. Sundaresan, J. Zhu, and J. P. Fernandes, "*Field Experiences with US OFDMA and Using US Profile Management", SCTE 2020*

[ANOMALY DETECTOR ICPHM 2020] J. Zhu, K. Sundaresan and J. Rupe, *"Proactive Network Maintenance using Fast, Accurate Anomaly Localization and Classification on 1-D Data Series", 2020 IEEE International Conference on Prognostics and Health Management (ICPHM), 2020, pp. 1-11, doi: 10.1109/ICPHM49022.2020.9187045.*

[C3 CableLabs] CableLabs Common code community, https://community.cablelabs.com/wiki/display/C3

