# Rethinking Customer Support

## Proactive Customer Engagement: Experimentation in Real-Time Data

A Technical Paper prepared for SCTE•ISBE by

**Joe Keller**
Executive Director Analytics
Cox Communications
6305 Peachtree Dunwoody Road, Atlanta, GA 30328
404-269-5938
Joseph.Keller@cox.com

**Sam Plant**
Analytics Engineer
Cox Communications
6305 Peachtree Dunwoody Road, Atlanta, GA 30328
404-269-1743
Sam.Plant2@cox.com

# Table of Contents

**Title** **Page Number**

Table of Contents .......... 2

Introduction .......... 3
1. What is Proactive Customer Engagement? .......... 3
2. Going Proactive .......... 4
3. Improved Outage Module Detection and Staging .......... 5
   3.1. The Opportunity Factory .......... 5
      3.1.1. Hypothesis Testing .......... 6
   3.2. The Experimentation Factory .......... 6
   3.3. Pilot Architecture .......... 7
   3.4. The Model Process .......... 8
   3.5. Tuning Model Triggers .......... 9
   3.6. Production Shadow Results .......... 11
4. What's Next .......... 11
   4.1. Post-Outage Offline Device Resolution .......... 12

Conclusion .......... 12

Abbreviations .......... 13

## List of Figures

**Title** **Page Number**

Figure 1 - Cross Industry Standard Practice for Data Management .......... 4
Figure 2 - CMTS Trap Example .......... 6
Figure 3 - Pilot Design Process .......... 7
Figure 4 - Pilot Architecture .......... 8
Figure 5 - Main Loop Process .......... 9
Figure 6 - Upper Control Limit .......... 10
Figure 7 - Second Potential Limit .......... 10
Figure 8 - Basement .......... 11

# Introduction

Multiple Service Operators (MSOs) traditionally interfaced with customers reactively. When experiencing service degradation, seeking education, or requesting new services, the customer's only real recourse is to contact their service provider and attempt to describe their needs to agents who often struggle to provide resolutions sourced from a multitude of applications, datasets, and data sources. In the past, attempts to proactively interface with customers were stymied by a comprehensive lack of data understanding, by low data velocity, and by the cost-prohibitive nature of the operational and technological capabilities required to identify issues before they impact customer.

...But the industry and technology have changed.

Proactive Customer Engagement (PCE) represents a cultural shift in how Cox Communications interacts with our customers. Leveraging probabilistic models, higher velocity data, and cloud-based technologies, Cox Communications seeks to shift customer interactions from a reactive to a proactive stance.

## 1. What is Proactive Customer Engagement?

Proactive Customer Engagement (PCE) seeks to address our industry's emerging number one challenge: improving customer experience while simultaneously driving down operational costs. Cox Analytics posited a solution: Predict and address customer needs **prior to** a customer contact, thereby saving customers the arduous task of explaining their issues to an agent, and simultaneously reducing the cost of serving these needs through traditional high-cost channels.

There are three engagement strategies for PCE:

- **Deflection**: Predicting a customer's intent and correctly addressing that intent upon receipt of a customer contact (e.g. improved outage detection, identifying remote pairing issues based on set top box errors, identifying the wrong HDMI input based on tuning and set top box error data, etc.).

  This initial phase of PCE serves as the cross-over point from reactive management of customer needs to proactive management by leveraging real-time information while interacting with customers via more traditional means.

- **No Outbound Contact:** Predicting a service degradation or out of service scenario that can be fixed without customer contact (e.g. device reboot, device re-authorization, device re-provision, firmware push, etc.).

  The second phase of PCE corrects issues that might otherwise manifest in a customer contact without the customer's knowledge. Fixes are only applied when the customer is not actively using their services.

- **Outbound Contact:** Predicting a service degradation or out of service scenario issue that cannot be fixed without customer contact (e.g. home technician visit, network technician visit, device swap, pairing remote, etc.).

 

This final phase of PCE combines customer behavioral data and customer service usage data to interact with customers according to their channel of preference (e.g. SMS, email, outbound dialer).

When engaging customers proactively, thorough consideration must be made for establishing the correct balance between 'being caring' and 'being creepy'. Initial experiments are designed for deflection scenarios with no outbound contact while ongoing analysis focuses on classifying the customer by preferred channel of interaction.

## 2. Going Proactive

The cable industry continues to strengthen its ability to react to customer needs by leveraging the once latent data produced by the network itself. Customer Premise Equipment (CPE) telemetry, network telemetry, and guided trouble-shooting platforms arm agents with the ability to collect data, trouble-shoot, and identify the best fix agent to address a customer's need.

To realize the vision of going proactive, PCE requires streaming/real-time data assets and advanced analytics capabilities to predict and address customer needs before the customer initiates contact. A non-trivial task.

To this end, it is necessary to determine the feasibility of proactive engagement. This is done leveraging a repeatable methodology: **The Analytics Lifecyle**.

Cox Analytics team employs the Cross-Industry Standard Process for Data Management (CRISP-DM) to manage PCE use cases through the **Analytics Lifecycle**. Hypotheses are developed and prioritized via the **Opportunity Factory**, promoted to pilot via the **Experimentation Factory**, and finally operationalized at scale via the **Deployment Factory**.

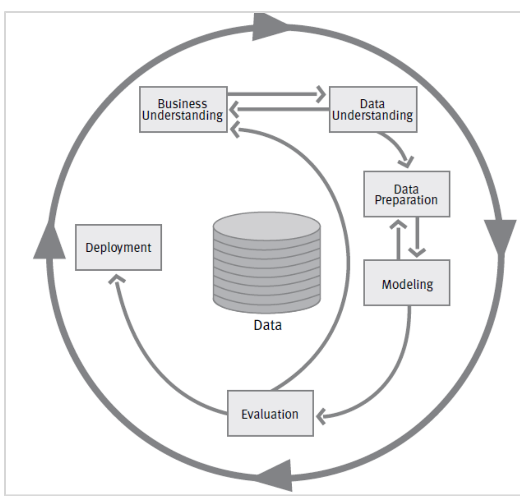


**Figure 1 - Cross Industry Standard Practice for Data Management**

**Opportunity Factory**

- Business Understanding: The process of developing an understanding of project objectives and requirements, translating this understanding into required data sets, and the performing preliminary data discovery and preparation.

 

- Data Understanding: Performing Initial data collection, insight development, and hypothesis development. (What data informs the objectives and outcome of each use case scenario? What is the quality, velocity, and availability of that data set?)
- Data Preparation: Definition of the logical data model inclusive of table, record, and attribute source selection and transformation (if applicable) and aggregation for consumption by operational models.
- Modeling: Design, build, and iterate on models while focusing on model recall, precision, and confidence. This includes shadow model training and the introduction/removal of features based on their ability to improve on model effectiveness.

**Experimentation Factory**

- Evaluation: The execution of a model in Pilot with continued iteration over model development and continued evaluation of model effectiveness. This will lead to a go/no go decision for enterprise deployment.

**Deployment Factory**

- Deployment: The enterprise-wide deployment of predictive/probabilistic models beginning with an initial pilot and extending to enterprise availability.
- Continuous Improvement: Continuous evaluation of the predictive/probabilistic models for effectiveness which includes reassessment of feature selection, further A/B testing, and channel-specific effectiveness.

The balance of this paper tracks our initial foray into proactive engagement through the Analytics Lifecycle, elaborating upon our initial use case: improved outage module detection and staging.

## 3. Improved Outage Module Detection and Staging

Project Off Ramp is a code-name for Cox Communication's improved outage detection and call deflection program. The program represents an initial foray into experimentation with proactive engagement, leveraging real-time streaming trap data to detect HFC network outages and stage IVR outage messages.

Historically, outage detection was based upon a combination of 1) the number customer calls received by the IVR within a given time frame for a common node, and 2) polling-based telemetry. These outage module detection algorithms lagged outages by 15 minutes or more. With the introduction of Off Ramp, we can use real-time, streaming trap data to rapidly detect an outage condition, allowing for the deflection of more calls, eliminating unnecessary truck rolls to the customer premise, expediting service restoration, and improving overall customer experience.

### 3.1. The Opportunity Factory

The first step on the path to proactive engagement begins with the discovery of available datasets that meet the dual requirements of being both available in real-time and containing leading indicators of customer issues. One such data source is real-time streaming CMTS Modem Online/Offline trap data that can be used to identify the status of a DOCSIS device.

We hypothesized that the streaming trap data could be used in combination with IVR messaging and work order blocking to reduce support calls and eliminate unnecessary truck rolls during an outage more quickly than the system that was in place.


```
"Identifier": "172.30.63.99  cdxCmtsCmOnOffNotification d4 04 cd d7 f4 47 0 3",
"Serial": "372440111",
"Node": "DT1XCAPC06",
"NodeAlias": "DT1XCAPC06",
"Manager": "MTTRAPD",
"Agent": "CISCO-DOCS-EXT-MIB",
"AlertGroup": "cdxCmtsCmOnOffNotification",
"AlertKey": "d4 04 cd d7 f4 47",
"Severity": 3,
"Summary": "Cable Modem Online(CMStatus: online; DownChannelIfIndex: 68617; UpChannelIfIndex: 396748)",
"FirstOccurrence": "2019-07-05T06:56:52",
"LastOccurrence": "2019-07-05T06:58:36",
"Type": 0,
"Tally": 2,
"ProbeHost": "FED1ISMV01",
"ProbeType": "MTTRAPD",
"Hostname": "
"MACAddress":
"IPAddress":
"Condition": "ONLINE",
"NEVendor": "CISCO",
"NEModel": "cBR8",
"NEType": "CCAP",
"SystemName": "SAN DIEGO"
```


**Figure 2 - CMTS Trap Example**

#### *3.1.1. Hypothesis Testing*

Can CMTS Online/Offline Traps Outage Detection Enable Us to Outperform Current In-Place Outage Detection?

This hypothesis above was tested using historical call and outage data combined with a POC model utilizing the real-time CMTS trap messages. Calls and truck rolls that occurred between the start of an outage detected using the new model but before the time the same outage was detected using the in-place model were counted toward the model's effectiveness.

This preliminary analysis confirmed the viability of the hypothesis. Even with allocations made for the subset of our customer base passing through the IVR during an outage that still want to speak with an agent, we saw a significant opportunity for additional call and truck deferrals. Table 1 represents annualized reductions in calls and trucks calculated during use case validation.

**Table 1 – Preliminary Business Case**

| | **Transaction Reduction Opportunity** |
|---|---|
| Technical Support Calls | 233,169 |
| Scheduled Truck Rolls | 38,815 |

With the business case validated, the use case proceeds onto model development and experimentation.

### 3.2. The Experimentation Factory

Experimentation begins with defining the conceptual architecture and process for tuning the outage detection model. In this case the streaming data source originates at the CMTS, NetCool collects these messages and forwards them on to a Kafka topic to minimize the latency of this pipeline. A VM running the outage detection model then consumes the messages.

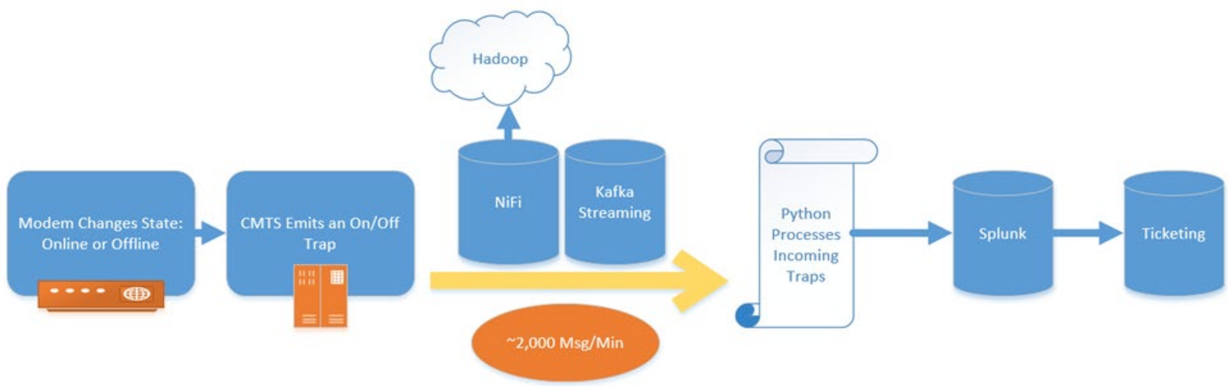


**Figure 3 - Pilot Design Process**

Within the model, messages are enriched with additional attributes that allow for the identification of a customer's node using the Upstream and Downstream interface card ID associated with the message. A state table is then updated with the online/offline state of the device.

The approach is analogous to methods of counting children in a classroom.

If you have counted the number of children in a classroom once, you can keep an accurate count of the children by increasing or decreasing your count based on how many pass in and out of the room rather than recounting everyone as you would in a polling system.

In much the same way, we populate an initial snapshot of online devices in the device state table and then maintain synchronization with any changes for added or removed devices. As an added precaution, daily polling updates remove devices not present for > 3 days from the device state table.

### 3.3. Pilot Architecture

A modular approach was chosen for the CMTS Traps outage detection model. This was done for three main reasons. First, we wanted to be able to expand the capabilities of the model without being forced to change the main model loop. Second, we needed the ability to make changes to data elements while the main outage detection model continued to function. This was specifically to enable maintenance on the device lists and to give access to node and device information in real time through an API. Lastly, we wanted to design a model that could easily be migrated to the cloud.

Our model relies on an in-memory database tool to manage a number of key data objects: modem mac address key value pairs to manage modem level state and location information, node key value pairs to manage aggregated state information for our HFC nodes, and a number of streams to manage communication between the main consumer loop and our governor process. Since we had a variety of data type needs, we selected Redis as our in-memory DB.

We designed this model to be easily migrated to the cloud in the future. The main module is essentially a message consumer. Using AWS Kinesis, Elasticache, API Gateway, and Lambda, we could replicate the architecture below while increasing the overall scalability of the model.

 

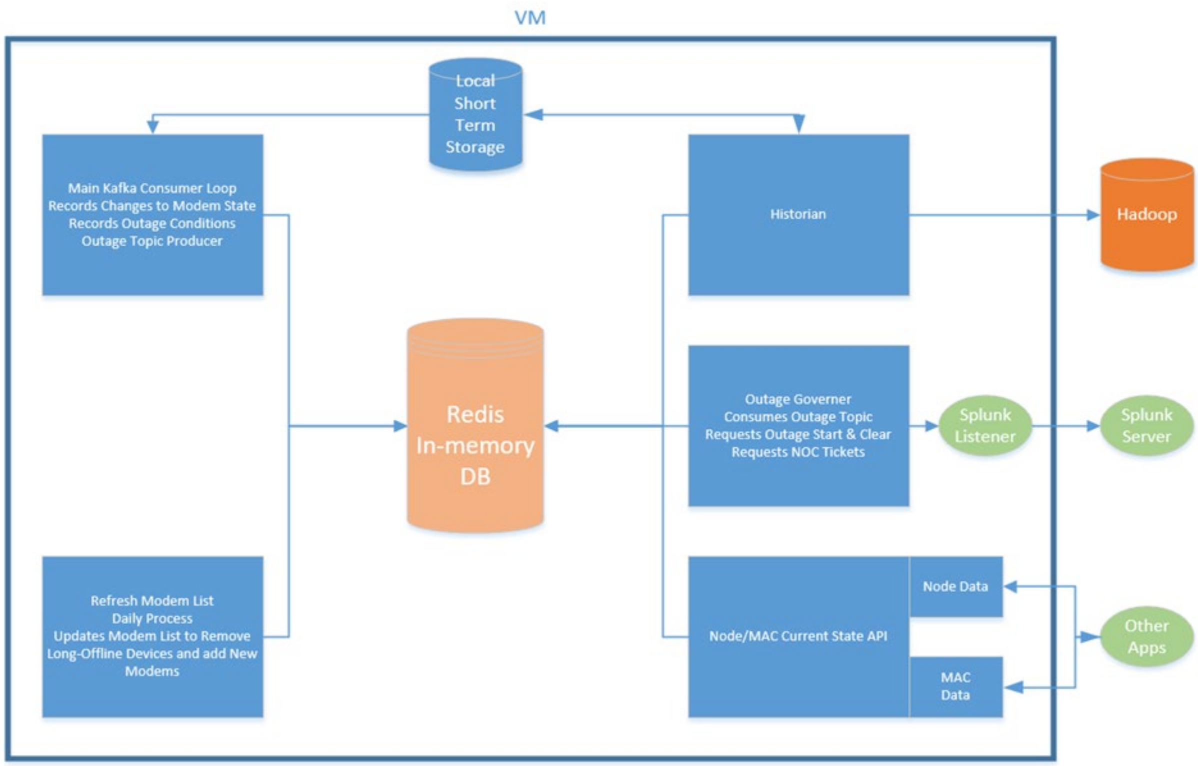


**Figure 4 - Pilot Architecture**

### 3.4. The Model Process

There are 4 processes that comprise the Off-Ramp model:

1. **Main Loop Process**: Core process to the model. The Main Loop process receives streaming trap data, applies attribution for mapping to billing system Site ID and node, updates the device state table, and checks for the Upper Control Limit for offline devices, thereby not taking any action, declaring an outage, or clearing an outage.
2. **Device List Update Process**: Process for maintaining an up-to-date device list. The Device List Update process leverages device polling data, comparing polled device lists to the device state table. Devices that have not been present in polling results for > 3 days are removed from the table. Devices that have not been present are added.
3. **Outage Trigger Governance Process**: Process that consumes outage messages from the Main Loop process. The Outage Trigger Governance process adds each message to a list of pending outage start/clear events per node.
    - If a clear is received after an outage and before 3 minutes, both the outage and the clear are deleted. (No outage is triggered.)
    - If an outage message is received after a clear and before 10 minutes, both the clear and the outage are deleted. (The outage remains active.)
    - If an outage is received and no clear is received before three minutes, the outage is triggered.
4. **Historian**: Leveraged for historical outage profiling and for weekly calibration of the Upper Control Limit logic. The Historian reads the outage stream and copies the data to a long-term

database while removing the messages from the stream. A database is used given that it makes the process of recalculating upper control limits easier.

Figure 6 below provides a detailed explanation of the Main Loop process.

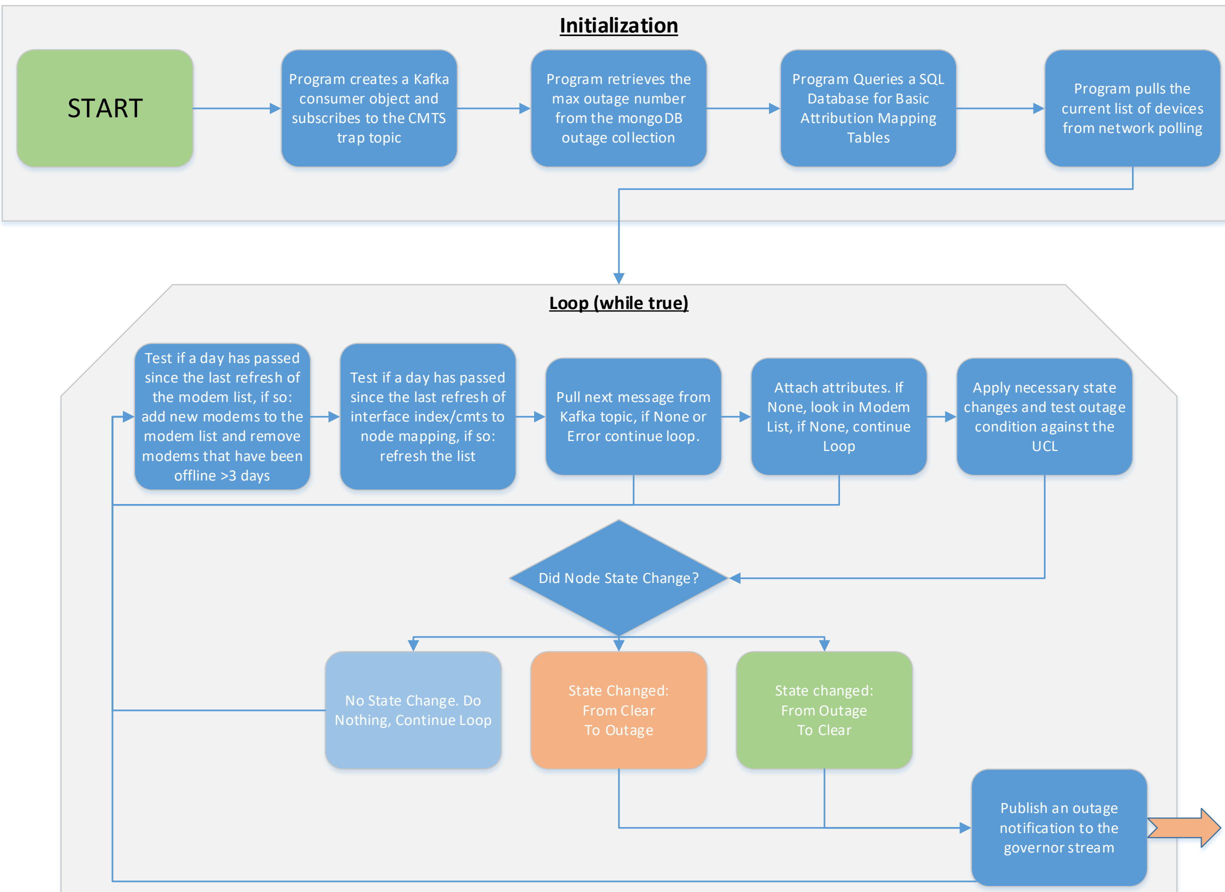


**Figure 5 - Main Loop Process**

## 3.5. Tuning Model Triggers

The most critical step in developing any model is calibrating that model for intended outcome. In the instance of the improved outage staging model, calibration focuses on the triggering logic for staging and clearing outages. Model tuning must account for small increases in offline modems due to customers restarting their devices without triggering an outage.

The upper control limit is the maximum of three potential values used to trigger the staging or clearing of an outage. If the number of offline devices exceeds the UCL for > 3 minutes, then an outage is triggered. If the number of offline devices for an active outage returns below UCL levels for > 10 minutes, then an outage is cleared.

1. The first limit is calculated using Median Absolute Moving Range (MAMR). This method utilizes the median absolute deviation methodology which is much more resilient when your data has outliers. Outlier events in terms of offline devices at an HFC node level represent a significant departure from the normal state of a node. While a healthy node usually only has a

few devices offline at any given time, and changes tend to occur in small single digit increments and decrements, outage events tend to occur suddenly and include dozens of users. This ruled out the use of standard deviation which tends to exaggerate outlier events and could result in an overly high upper control limit (especially if there is a history of whole-node outages).

As indicated in figure below, the center line represents the median number of offline devices by node. We add to this our sigma value multiplied by 3. Sigma is calculated as 1.0483 times the median of the absolute value of the differences (over time series) in offline devices between measurements by distinct node.

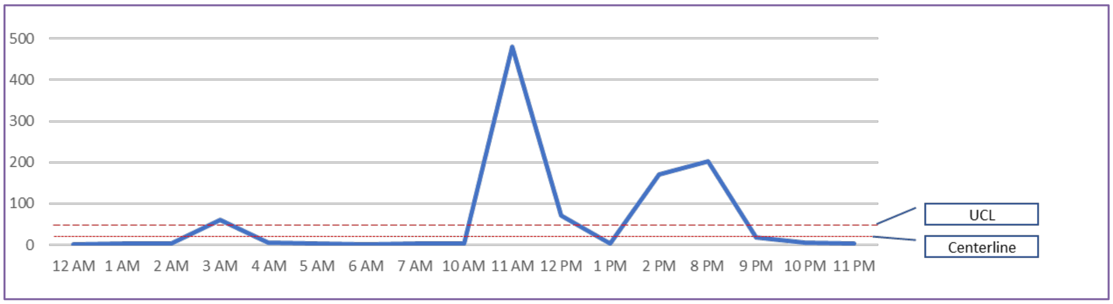


**Figure 6 - Upper Control Limit**

2. The second potential limit is calculated as a percent of the total number of devices on a node. As indicated in figure n.n below, if the number of offline devices meets or exceeds 2.5% of population of devices on a node, an outage will be triggered. This was chosen to avoid over sensitivity on lower activity nodes.

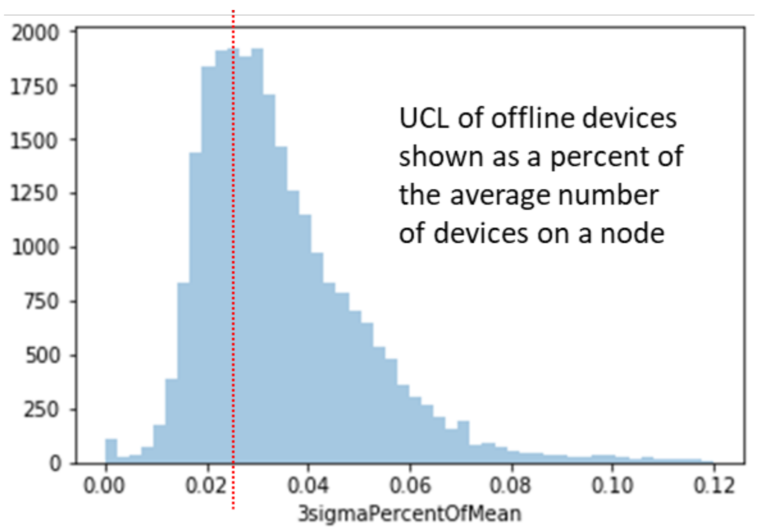


**Figure 7 - Second Potential Limit**

3. The 'basement' or minimum number of offline devices required to trigger an outage is established leveraging the actual value distribution of offline devices on a node. As indicated in figure n.n. below, 5 offline devices is the minimum number of devices allowed to trigger an outage from any node.

 

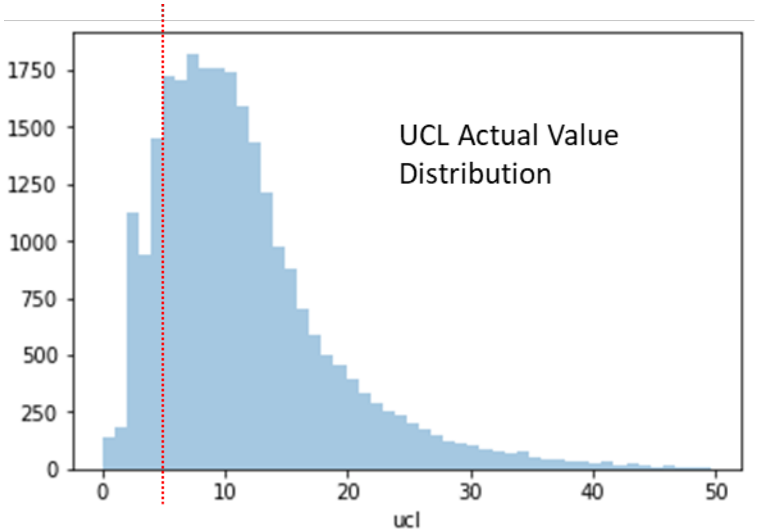


**Figure 8 - Basement**

### 3.6. Production Shadow Results

Production Shadow Models leverage production data to validate the effectiveness of a model without invoking action. In this instance, no outages were declared; however, the times that the model triggered outages are logged and compared against instances of observed production outages. The corresponding volume of calls and trucks that could have been avoided is calculated based upon the outage start time according to the Off-Ramp model vs. the outage declaration time according to the legacy model.

The results represent a significant improvement in outage deflection based upon the improved outage module's detection capabilities leveraging streaming trap data.

**Table 2 – Production Shadow Results**

| | Shadow Run | Transaction Reduction Opportunity |
|---|---|---|
| Technical Support Calls | 1 | 140,950 |
| Scheduled Truck Rolls | | 27,314 |
| Technical Support Calls | 2 | 176,428 |
| Scheduled Truck Rolls | | 32,219 |

Based upon these results the Off-Ramp outage detection and staging model will graduate to production pilot over the course of Q3 2019.

## 4. What's Next

As mentioned above, the Off-Ramp outage module detection and staging experiment represents an initial foray into proactive experimentation. That said, there are several use cases currently undergoing feasibility assessment that represent the evolution of our vision for proactive customer engagement. Each use case considered for PCE must meet the minimum qualifications of being enabled via real-time datasets that serve as a leading indicator for customer issues.

The following is an example of a use case that is currently undergoing discovery and validation as next priority for PCE.

### 4.1. Post-Outage Offline Device Resolution

As a fast-follower to improved outage module detection and staging, teams are evaluating the feasibility of leveraging the same streaming CMTS trap data to identify customer devices that have not returned to an active online state after an outage. The post-outage offline device resolution experiment focuses on the 75th percentile of devices that remain offline following an all clear.

By detecting post-outage offline devices, we can proactively engage customers to restore their service to a healthy state, eliminating calls and improving customer experience.

# Conclusion

Proactive Customer Engagement represents a new frontier in customer service. By leveraging increasingly available real-time datasets and harnessing the burstable compute power of next generation analytics frameworks, the cable industry has an opportunity to realize a cultural shift in customer engagement, pivoting from a reactive to a proactive stance. The crawl, walk, run approach to realizing PCE begins at familiarization with real-time data-sets and applying them to cross-over use cases such as outage deflection.

As our results demonstrate, there are considerable cost savings to be realized by harnessing the power of these real-time datasets.

1. We observed an improvement of ~ 160,000 calls and ~ 30,000 truck rolls through early outage detection in combination with IVR messaging and truck roll work order blocking.
2. We validated the application of a real-time dataset to a high-performance model leveraging next generation technology that is readily transferrable to the cloud.
3. We identified a fast follower use case for identifying those devices which have not returned to an online status after an outage is cleared for proactive remediation and customer engagement.

It is through the practical application of real-time datasets to achievable use cases that we will realize our evolution into proactive customer engagement.

 

# Abbreviations

| | |
|---|---|
| API | application programming interface |
| AWS | amazon web services |
| CMTS | cable modem termination system |
| CPE | customer premise equipment |
| DOCSIS | data over cable service interface specification |
| HDMI | High definition multimedia interface |
| HFC | hybrid fiber coax |
| IVR | Interactive voice response system |
| MAMR | median absolute moving range |
| PCE | proactive customer engagement |
| UCL | upper control limit |

 