AUTOMATIC ROOT CAUSE ANALYSIS IN MOBILE NETWORKS

Julio Arpírez
CTO @ Zhilabs
jarpirez@zhilabs.com
• The Problem
• The State
• The Analysis
• The Conclusion
THE PROBLEM

• Today’s mobile networks should ideally
  • **Be fully reliable and accessible** (ideally 100% of the time)
  • Be usable for a wide variety of services (e.g. CS Voice, SMS, PS, VoLTE)
  • Serve subscribers and machines through different technologies (2G/3G/4G) and devices
  • Provide perfect communications for person-to-person, person-to-server (e.g. OTT services), M2M and IoT

• A network is a very complex system and consequently problems arisen are often difficult to diagnose and correct
  • All of us know the WAR-ROOM to WAR-ROOM game

• Problems in the network have a **huge direct financial impact** (revenue loss and the resources associated to operate and troubleshoot the network) as well as **indirect** (loss of credibility and impact on the brand)
THE STATE

Network State (Huge Dimensionality)

Problematic Operation Universe → Root Cause Analysis and Fixing

© Zhilabs 2015
THE STATE
SOME NUMBERS

50000 e/nodeB x 100 RNC x 20 SGSN/MME x 10 GGSN/PGW = 1000 Million Service Paths
10 State Variables x 1000 Million = 10000 Million Options to Explore
THE STATE
ADD MORE ITEMS...

Network Related
• Technology Type (2G/3G/4G)
• Policy servers
• OCS Systems
• DPI Nodes
• Optimizers
• ... 

Device & OTTs
• Device types & models (typically a small network has around 10000 active models)
• Internet servers providing fancy services (e.g. Google, Facebook, YouTube, Periscope, Twitter)
• OTT Apps (e.g. WhatsApp, SnapChat)

Keep on multiplying… The number of combinations to explore in quest of failures is unmanageable

© Zhilabs 2015
Given the state, we need to identify the boundaries of the Tolerable Operation Universe. Typically, this Tolerable Operation Universe is delimited by a Baseline for certain KQIs/KPIs (dimensionality reduction) – Typical KQIs include accessibility, retainability, speed, etc. + Confidence Interval. The Confidence Interval may be Hard – SLAs (typically, human-defined) or Soft – Automatic Baseline Computation. Deviations from the baseline + confidence interval are usually due to problems (deviations from normal behavior).
ROOT CAUSE ANALYSIS
THE FLOW

Real-time continuous production and analysis of all the indicators

Proactive Analysis

Detection of a deviation triggers multi-dimensional search

Multi-dimensional Search

An interactive RCA report is produced detailing all the steps followed and a ranking of the most probable Root Cause

Interactive RCA Report

An Engineer can use the report to understand the pursued steps and perform further investigation/ validation via additional drill-downs. It can also provision new rules/thresholds in the engine

(Optionally) Additional drill-down

Machine

Engine

© Zhilabs 2015
• Indicators are produced for all the dimensions
• A search of the values at the different dimensions allows to isolate origin of the deviations

© Zhilabs 2015
• The same algorithms available in FlowSight are used by the engine under the covers:
  • Forecasting: ARIMA
  • Anomalies: cyclostationary analysis, AR
• AR models are good at detecting anomalies
• ARIMA Methods take into consideration seasons and trends

Upper & lower thresholds are automatically calculated, based on the typical deviation around the baseline

Typically narrower for high values and broader for low ones
The system computes automatically Distribution Functions (PDF and CDF).
PROBABILITY DISTRIBUTION FUNCTIONS:
OUTLIERS

Long tails are explored by the automatic RCA
EXPERT RULES: FORMULAS

Clicking over a KPI shows the corresponding formula. The Formula Editor allows an operator to flexibly define complex KPIs.
1. **Proactive Analysis:**
   - Input: Nodes, Service Path, Devices, OTT
   - Additional inputs (multi-facet analysis):
     - Operational logs / Work logs (detailing changes performed in the network)
     - Performance Mgmt / Fault Mgmt / Logs from the different nodes
   - Multi-dimensional Indicator production
   - Four elements:
     - Anomaly detection (autoregressive and cyclostationary algorithms)
     - Thresholds
     - PDF outliers
     - Expert Rules

2. **Multi-dimensional Search:**
   - Multi-dimensional analysis of deviation to locate commonalities
   - Determination of most probable root-cause (ranking process)

3. **Interactive RCA report:**
   - Details steps (1) and (2) via clickable reports
   - Allows an engineer to validate the RCA and/or further investigate from the machine-produced baseline
1. Pattern similarity:
   • Uses time-range/indicator that signaled the anomaly as an input
   • Computes similarity for that scope in other dimensions
   • Correlation used as a similarity function between the time series

2. Ranking:
   • Dimensions effectively become a loosely-coupled service graph graded by the similarity function
   • A variation of Personalized PageRank algorithm is applied:
     • Random walk
     • Similarity score used to modulate transitions and teleportation

Intuitively equivalent to several Engineers drilling-down through the different dimensions driven by similarities in the time series.
THE ANALYSIS

Automatic Root Cause Analysis Demos
AUTOMATIC ANOMALY DETECTION
APPLE SERVICES DOWN
RCA: USE CASES

INTRODUCTION
Anomalies are presented in the main WALL by Anomaly Detector, showing:
- Anomaly Level (e.g. Network)
- Indicator (e.g. HTTP Transactions)
- Anomaly Date (e.g. at 20140222194500)
When accessing to the anomaly report, a timeline is shown for a longer period rather than only the anomaly itself in order to check the signal evolution of the presented indicator.
By clicking the ‘anomaly flag’, a popup panel appears listing all anomalies for the selected period and also allowing to perform the Root Cause Analysis of this anomaly.
RCA NETWORK ANOMALY IN PDP CONTEXT ACTIVATION FAILURE RATIO INDICATOR

In this case, RAT-TYPE = GERAN shows the highest correlation. Root cause analysis is performed and system shows as a dashboard the FACTS and the Possible Issues found for the anomaly.

System drills-downs on GERAN for the APN dimensions and finds issues. And APNs found by RAT are presented in a set of ranked results.
An anomaly is found at Network Level in HTTP transactions indicator
Network Anomaly in HTTP Transactions Indicator

© Zhilabs 2015
**NETWORK ANOMALY IN HTTP TRANSACTIONS INDICATOR**

It is possible to see how volume for WhatsApp and Viber increase during the anomaly period.

It seems to be WhatsApp, suffering a system downtime during the anomaly period.

© ZhiLabs 2015
An anomaly is found at Network Level in Time To Stream Start Indicator

© Zhilabs 2015
RCA NETWORK ANOMALY IN TIME-TO-STREAM-START INDICATOR

First fact founds that issue happens for all GGSN nodes.

One site has high correlation with the issue. Top 10 are shown.

© Zhilabs 2015
An anomaly is found at Network Level in RTT-Client Indicator
Second fact showcases the issue for RAT=MIXED value

RCA creates a third fact clearly showing that Device= Tablet and APN=PREPAID is the main cause for this anomaly
THE CONCLUSIONS

• Root Cause Analysis can be seen as a search in a huge multi-dimensional space
• Different approaches are possible
  • Human experience (e.g. rule-based systems)
  • Fully automatic root-cause analysis (ML-aided systems, like time-series correlation techniques)
  • Mixed
• Once fully automatic tools are in place, the network can become self-repairing and adaptive, realizing the vision of self-organizing networks
• Root Cause Analysis applications do not only cover the network but also individual subscribers (e.g. Customer Care where call-dispatch time should be as short as possible)
THE CONCLUSIONS TOWARDS THE FUTURE

• Machine-Learning to the rescue!
  • To speed up troubleshooting
  • To detect offending elements before they become a real problem
• Close the Loop in the Network automatically
  • To fix problems
  • And also, to make your network a Chameleon! The so-called Self Organizing Network (SON)!
  • The network will adapt to the changing conditions and adapt itself by using the required resources (NFV)!

© Zhilabs 2015
THANK YOU!