

AI-based Root Cause Analysis of CPD (high noise) in DOCSIS networks

Georg Heiler | 2022-05 ANGACOM

Magenta



COMPLEXITY
SCIENCE
HUB
VIENNA



Problem & search process description for upstream high noise:

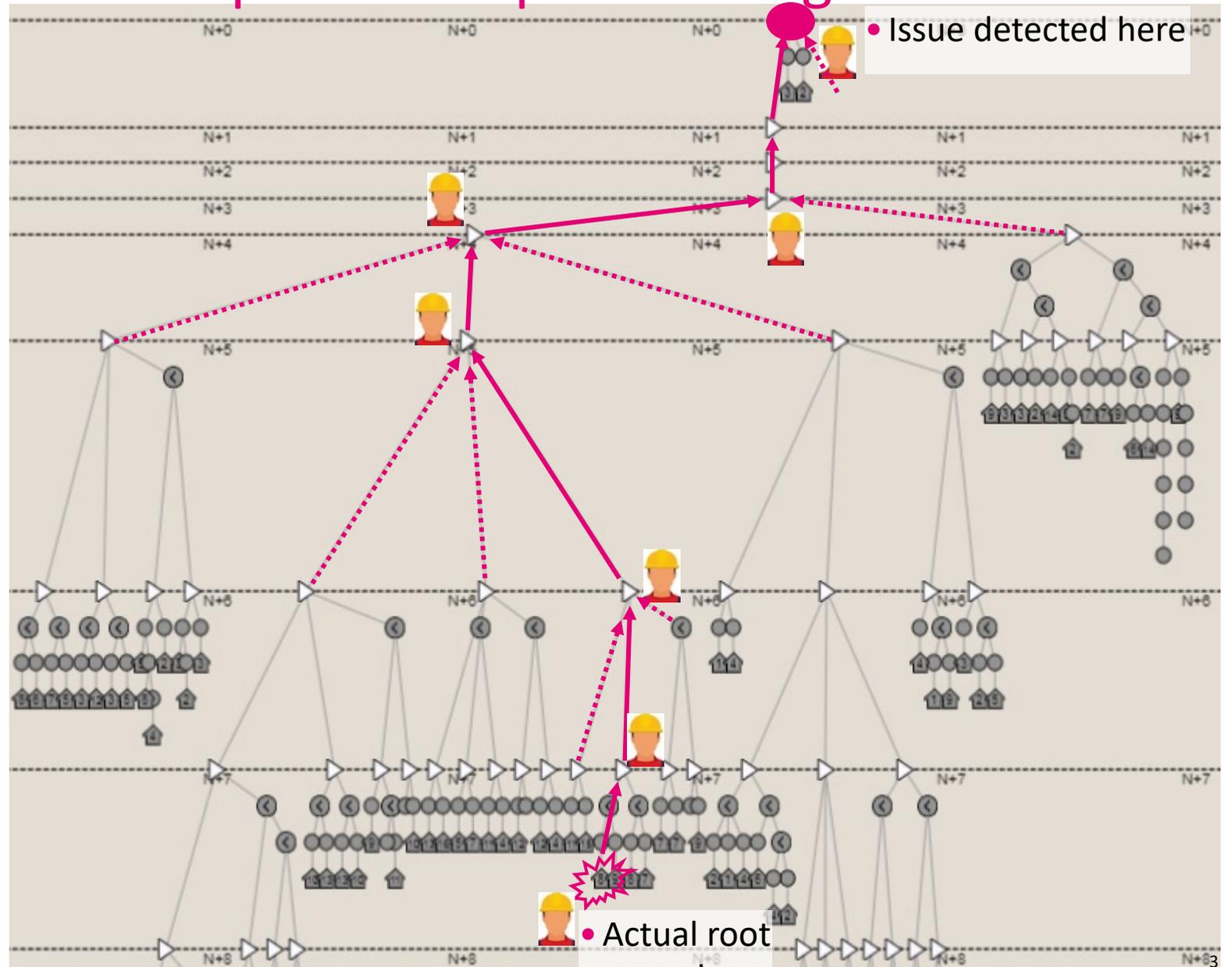
- **Problem:**

- Issue detected and root cause is unclear
- Many customers are affected as the disturbance is transmitted down to the root
- Identification of root cause is slow and manual

- **Customer impact:** Often service no longer usable

- **Task:** For a given incident direct the technician to the address of the root cause

- Reduce time spent for a ticket
- Reduce downtime and increase customer satisfaction

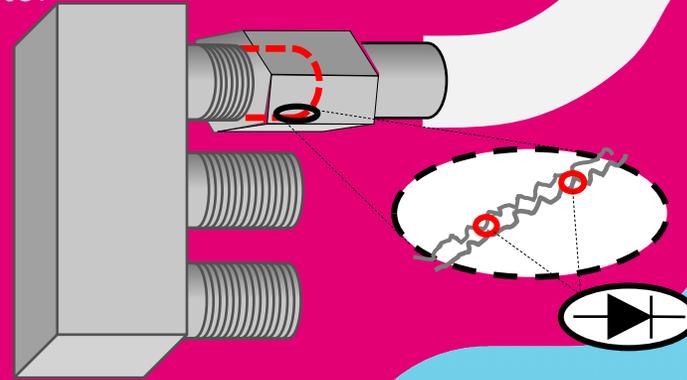


Business rule

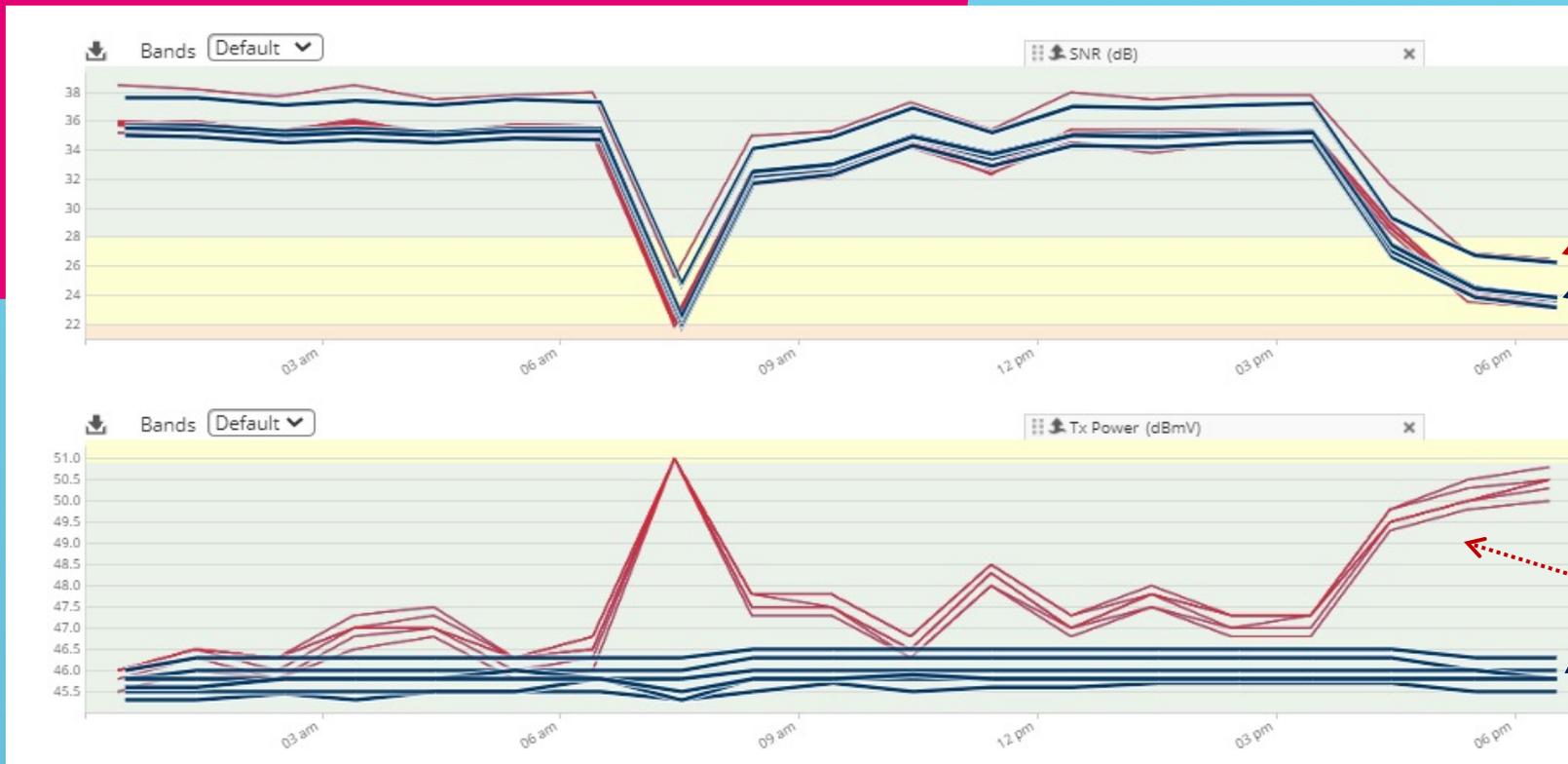
Largest TX change just before incident

Magenta®

incorrectly or loosely installed connector



Corroded connections, connectors not tightened.



Upstream-SNR

Single Modem

Network AVG

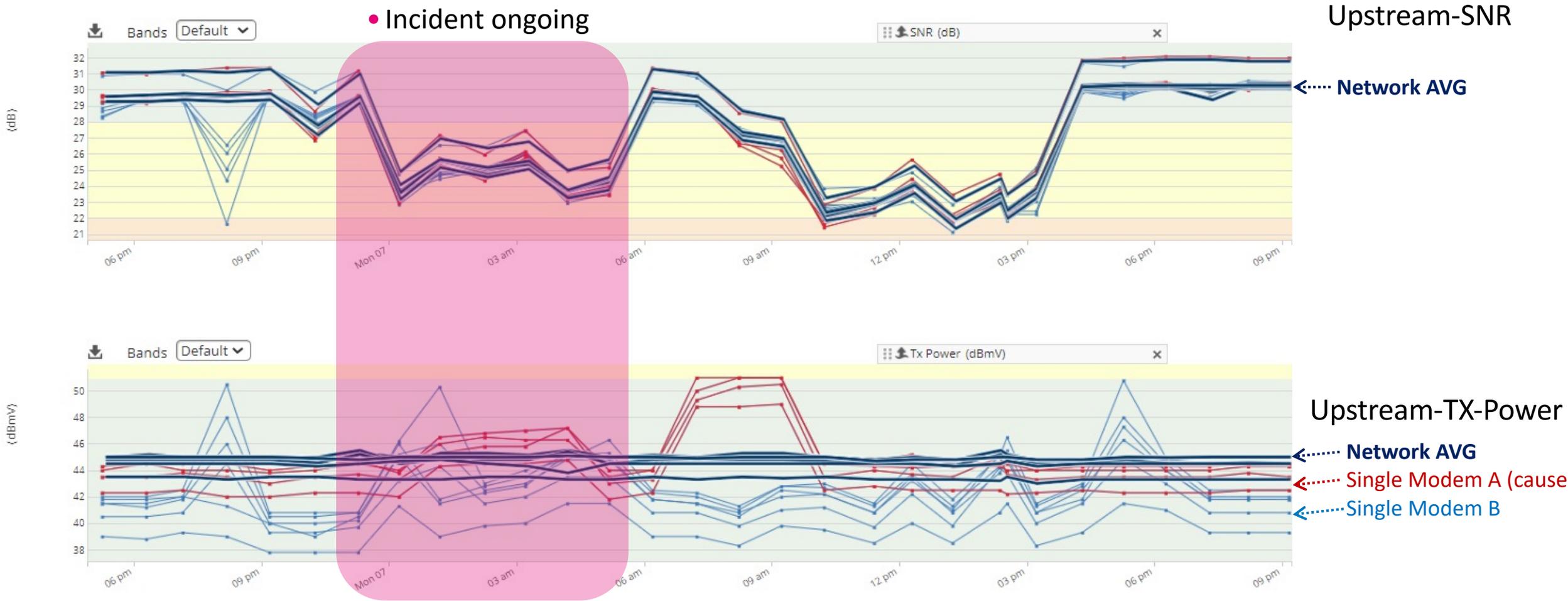
Upstream-TX-Power

Single Modem

Network AVG

Ambiguity of business rule

• Incident ongoing



Solution approaches: Levels of proactivness

01

Business rule

- Simple, well understood by technicians
- Device with largest TX change shortly before incident
- Can fail in ambiguous scenarios

02

AI assistance

- Works well also in complex scenarios
- Steer technicians directly to problematic device.
- Learn from the data & feedback of technicians (i.e. easy addition of new features)

03

AI insurance

- Predict incident and work proactively on its resolution
- Evaluation of quality (fixing non-existing problems)?

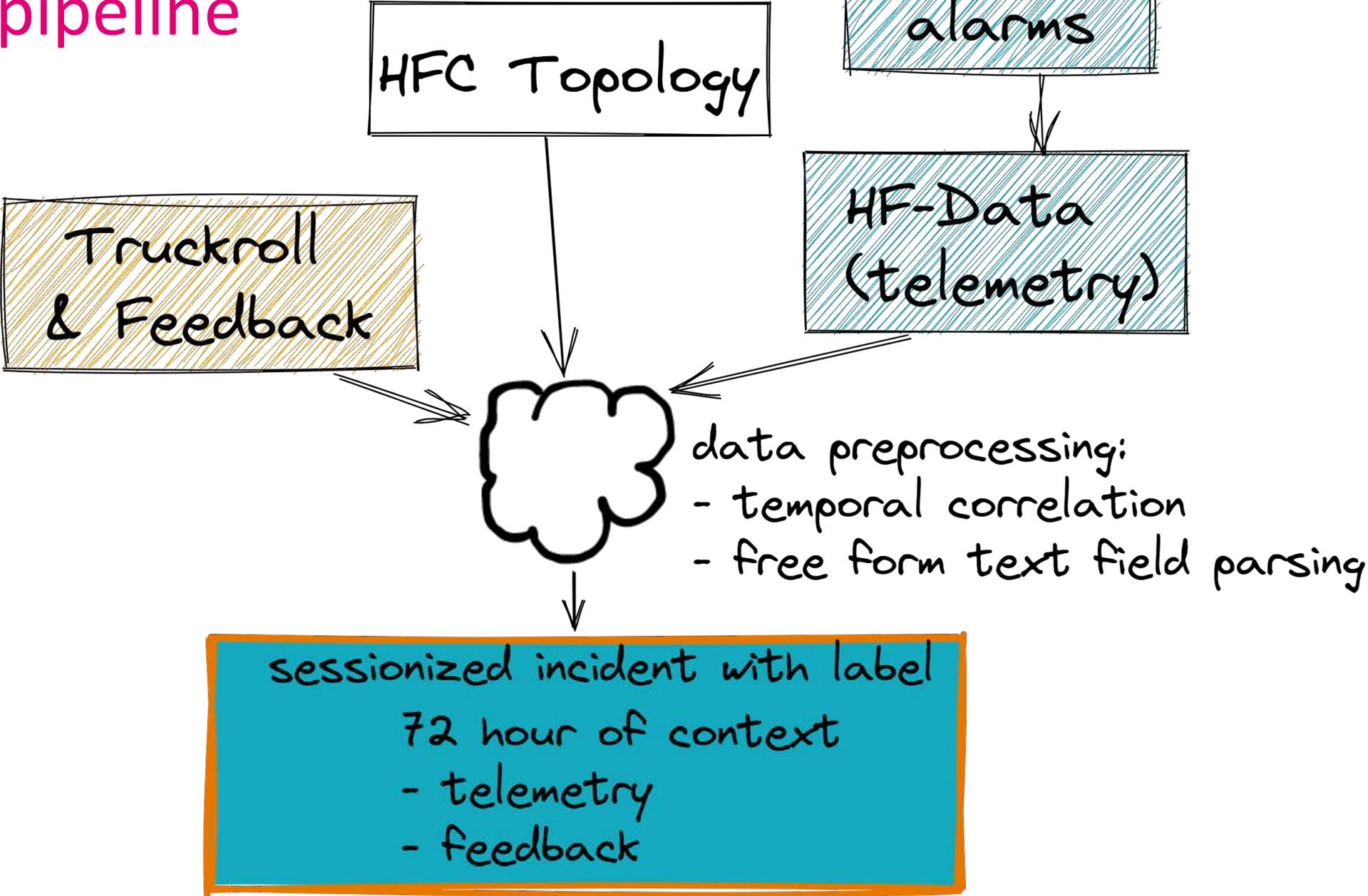


Improve network quality

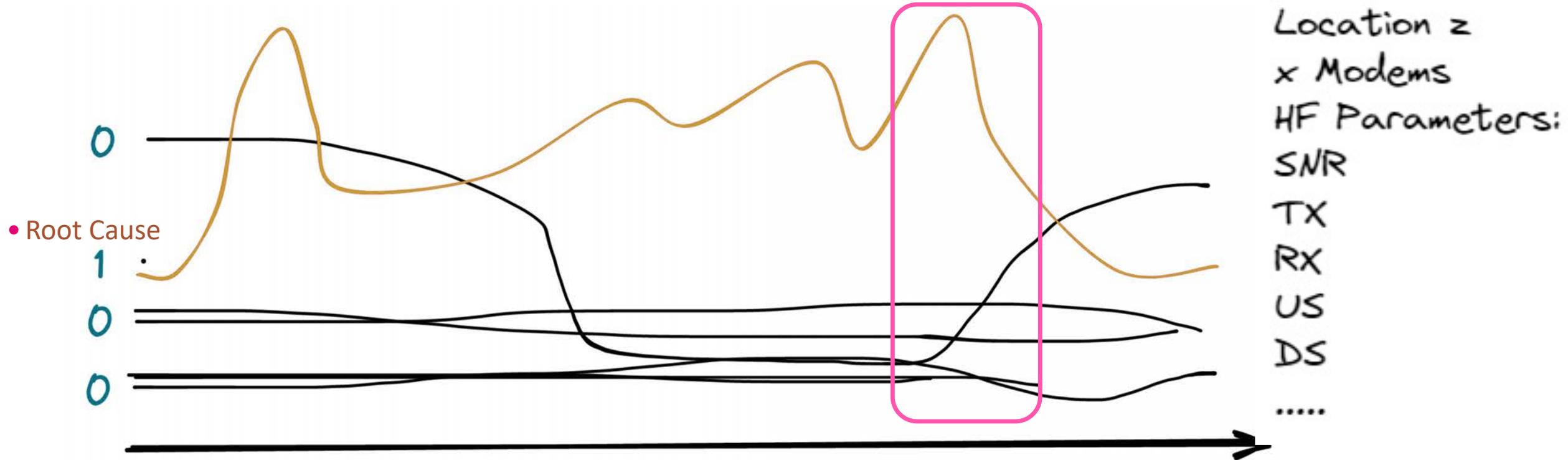
Increase customer satisfaction by resolving incidents faster:

- Reduce time customers are not satisfied with the service. Ideally fix the problem before they notice it
- Reduce incident resolution time (spent by our field technicians)

Data pipeline



Context of incident session window



session window with 72 hours
 for all devices in a fiber node area
 for all metrics (SNR, TX, RX, CER, CCER, Mreflection)
 for both up- & downstream

How does the Machine Learning pipeline work?

Sessionized Incidents with Labels

- Input: Data
- Feedback: Desired result

Training of Machine Learning Classifier

- ML model can automatically learn rules based on the data and desired result
- **Baseline: Business rule**
- Traditional models
- Gradient boosted trees
- Deep learning

Ranking of predictions [0.0 -> 1.0] consider top-k

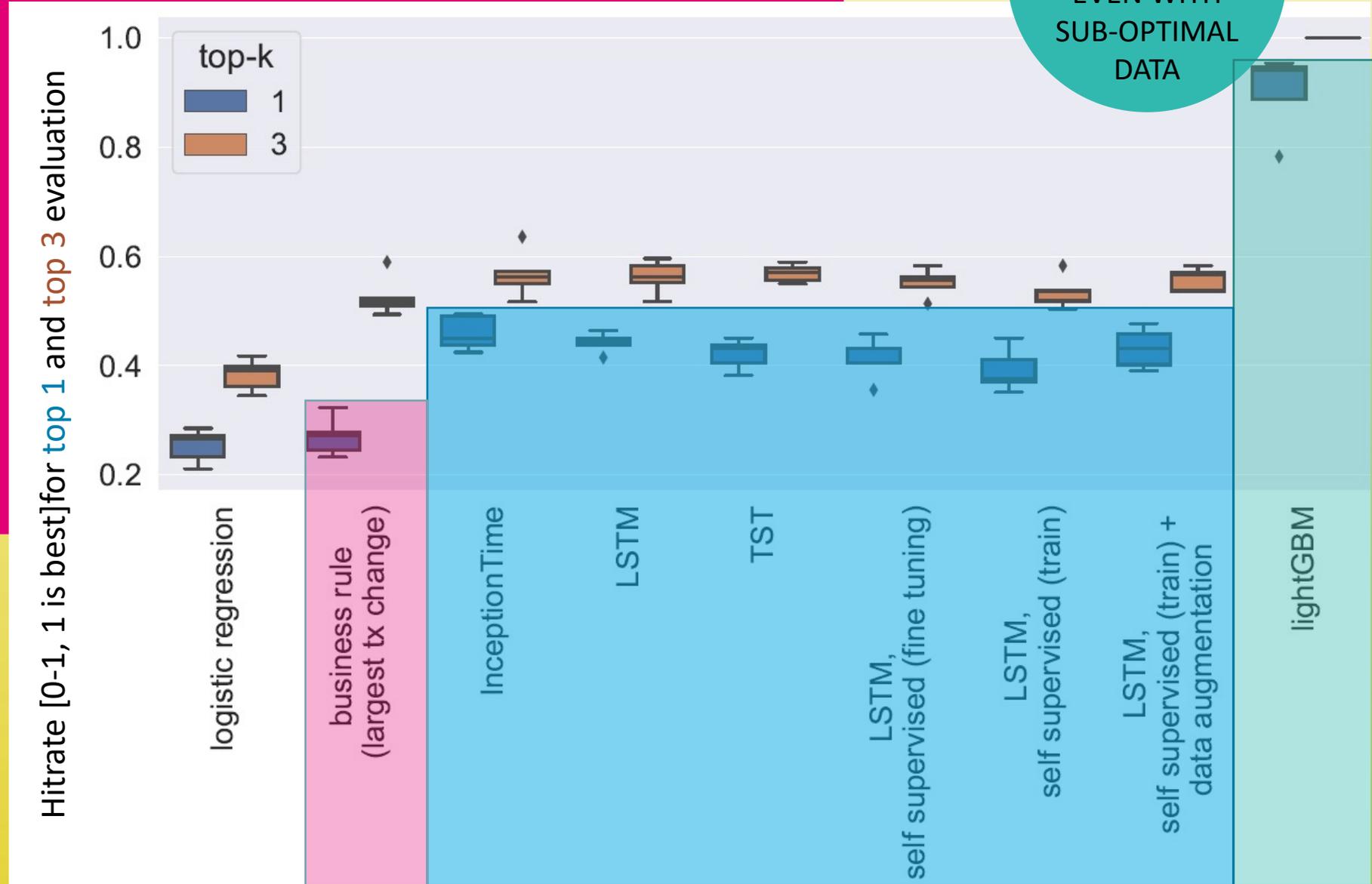
- Each network segment and incident is different
- Dynamics of incident automatically considered (Softmax)

AI assistance

2.3X better top-1 hit rate than business rule

Magenta®

- Baseline: Business rule
- Traditional models
- Gradient boosted trees
- Deep learning



AI insurance

Rethinking operations processes – a first step: Looking into the future
Scientific experimentation ongoing

Findings

01

Data Quality

- Build E2E feedback loops
- Quality & Usability: Use drop-down field in app for FF technicians to avoid incorrect (misspelled) amplifiers

02

OPS integration/streaming

- **React to events** and not conduct error-prone data archaeology (reconstructing the context)
- Central streaming ledger for real-time integration of tools of various vendors would be needed

03

Improve classifier

- Use AI instead of business rule: Prototype indicates **2.3x** improvement
- Experiment with forecasting
- ST clustering for flaky incidents (heatmap)



Improve network quality

Any AI usecase will profit from this solid foundation.

