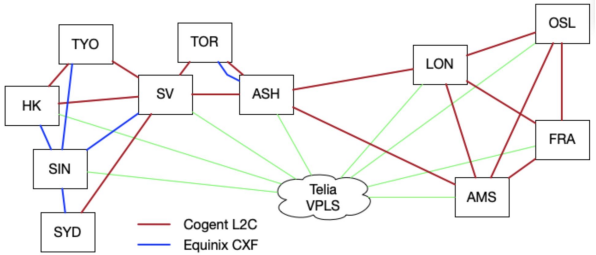
Crosslayer Network Outage Classification Using Machine Learning

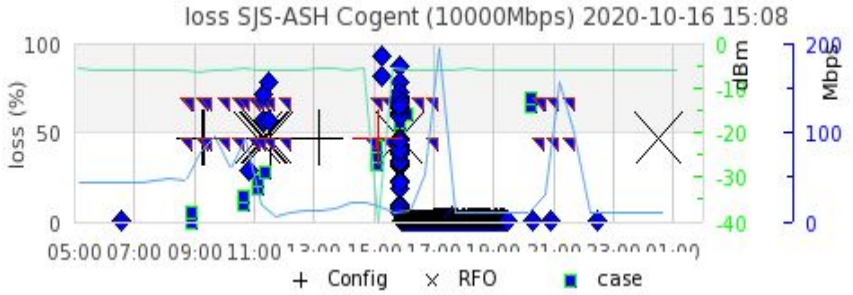
Jan Marius Evang
Azza H. Ahmed
Ahmed Elmokashfi
Haakon Bryhni

Simula Metropolitan Center for
Digital Engineering

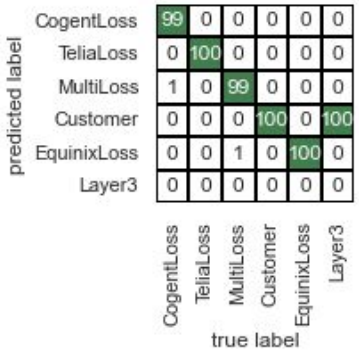
This paper focuses on the real-world challenges of classifying network outages



Testbed and problem statement

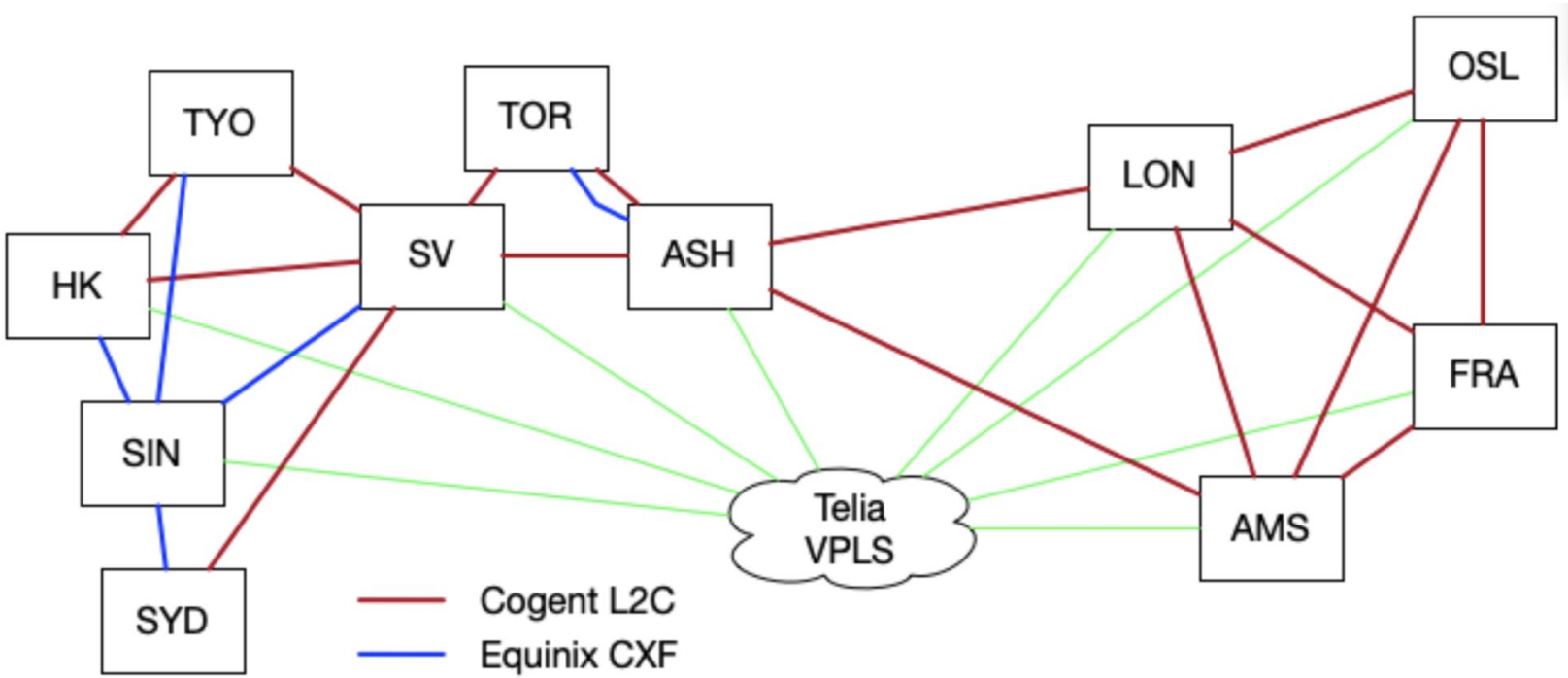


Methods and data



Results and discussion

A global network for quality-aware network traffic



Network problems are frequent

2 years of data
2019-08-08 to 2021-10-01

717352 Packet loss incidents
(Well within SLA)

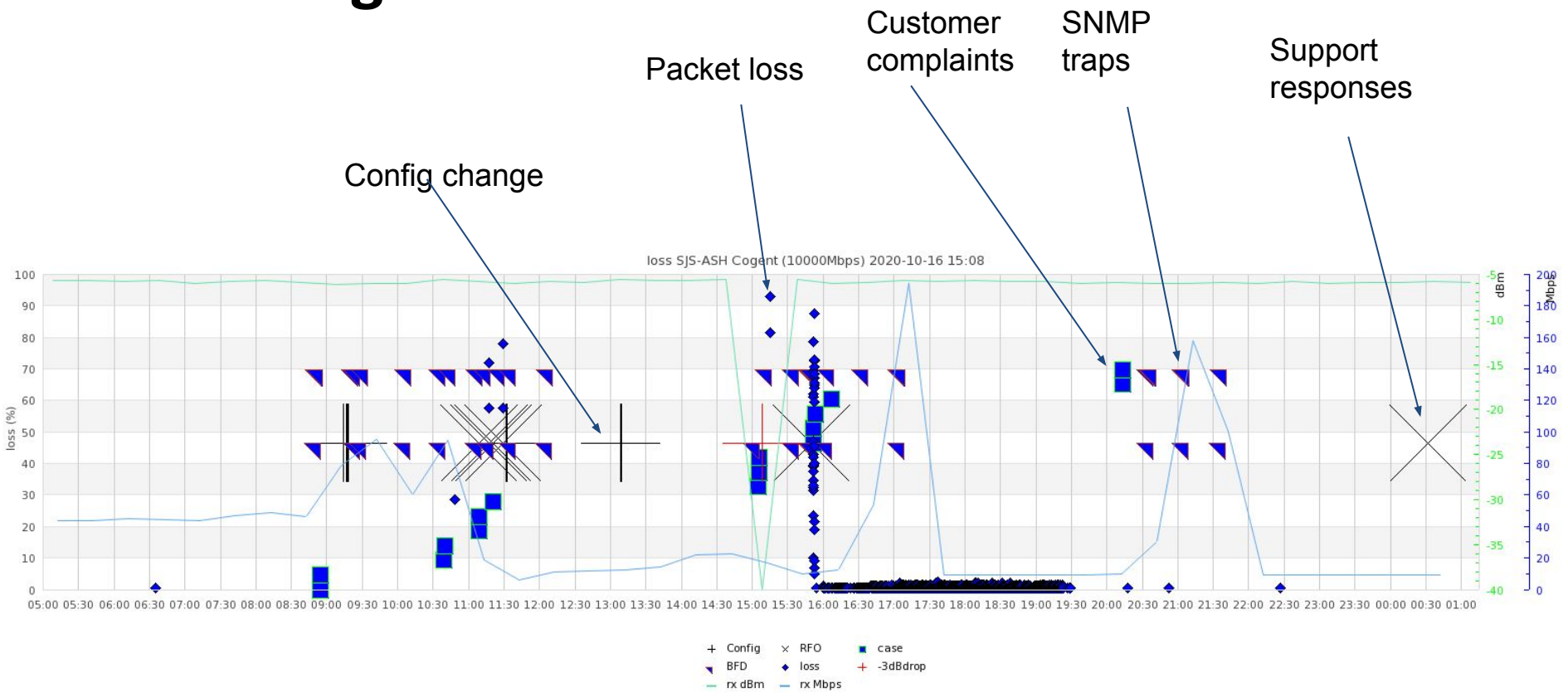
2855 Customer support cases
1014 Received a proper RFO (!)

Data collected

- Optical Signal strength
- Interface error counters
- Buffer overflow counters
- Software crash logs
- Config change logs
- BFD SNMP traps
- Layer2 packet loss (UDP-ping)
- Customer complaints data
- Customer support responses

Customer service tools to visualize all data

→ Confusing and ineffixient



In-depth analysis of causes of complaints

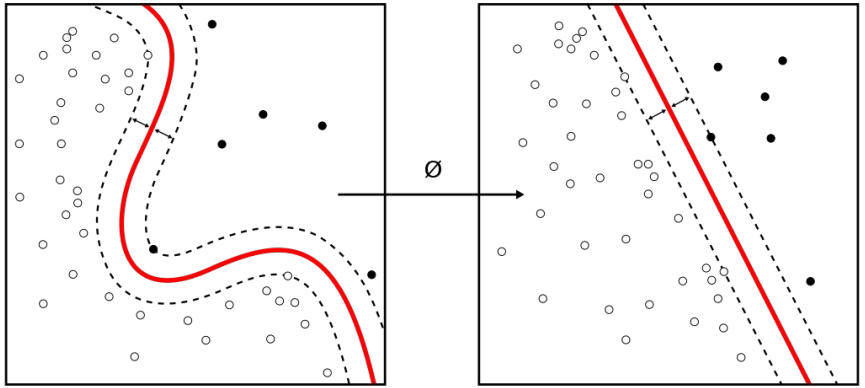
Also: Identify which measurement data is likely to be useful

Table 1: Manually classified causes.

Class	Cause	Count
MultiLoss	Multiple Layer2 providers	870
CogentLoss	Cogent's network	556
Customer	Customer's equipment	411
TeliaLoss	Telia's network	316
Layer3	Layer3 only	222
InternMaint	Internal maintenance	114
Optic	3dB Optical change	74
ProvMaint	Provider maintenance	70
EquinixLoss	Equinix Cloud Fabric	58
SubseaCable	Subsea cable outages	42
EquipFail	Equipment failure	40
FiberCut	Fiber cut in provider network	39
Layer1	Leased Layer1 lines	18
Metro	Metropolitan area links	18
DoS	Denial of Service attacks	4

Machine Learning - a very brief introduction

Basic goal: Divide a set of data points into two groups.



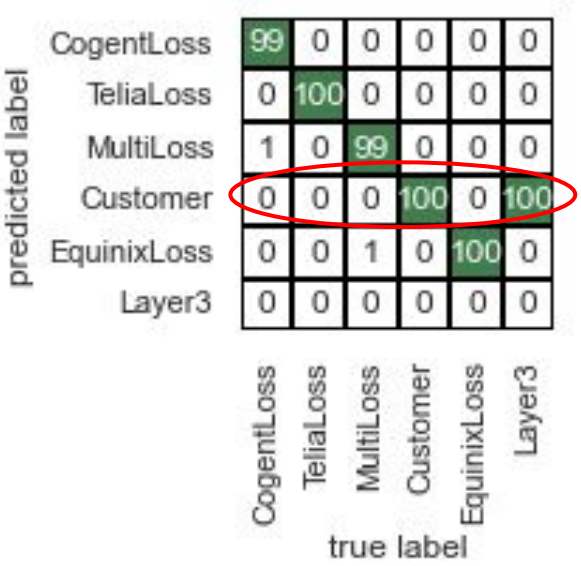
Support Vector Machine (SVM): Transform data such that the Decision Boundary is as wide as possible
→ Better at classifying “new data”. Less susceptible to overfitting.

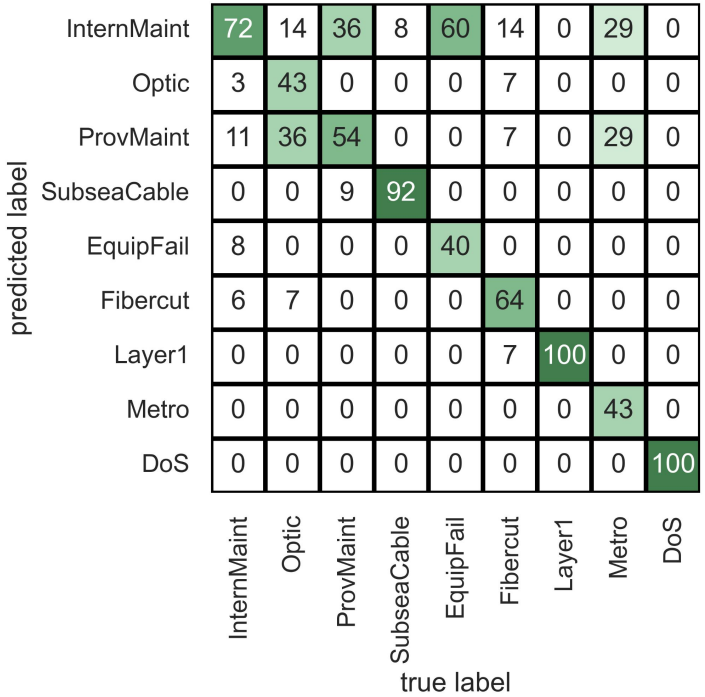
Results show that a two-stage machine learning process is optimal

Stage 1:
Identify the most common outage causes by using BFD data only.

Stage 2:
Identify other causes by using BFD + UDP-ping.

Classification accuracy





BFD only:
Very high accuracy for the most common causes

BFD + UDP-ping:
(for the Layer3/Customer class)
Reasonable accuracy for the remaining cases

In summary,.....

Machine learning proved to be well suited to classify outages.

Using BFD SNMP traps to log Layer2 events was very useful.

UDP packet measurements were necessary if we wanted to classify problems in other layers.

simulamet

Questions?

OSLOMET