\Orchestrating a brighter world NEC

[ITU-ML5G-PS-036]

Radio Link Failure Prediction

Dheeraj Kotagiri, Anan Sawabe, Takanori Iwai
System Platform Laboratories
NEC Corporation, Japan

Problem Description

Background

- Wireless backhaul links are low cost alternative to connect core network with the subnetworks.
- 5G communication enforces high reliability requirements (99.999%) on backhaul links as well.
- Reliable backhaul links are NECESSARY component to achieve 5G communication era.
- They typically use mmWave which are inherently susceptible to adverse weather conditions.

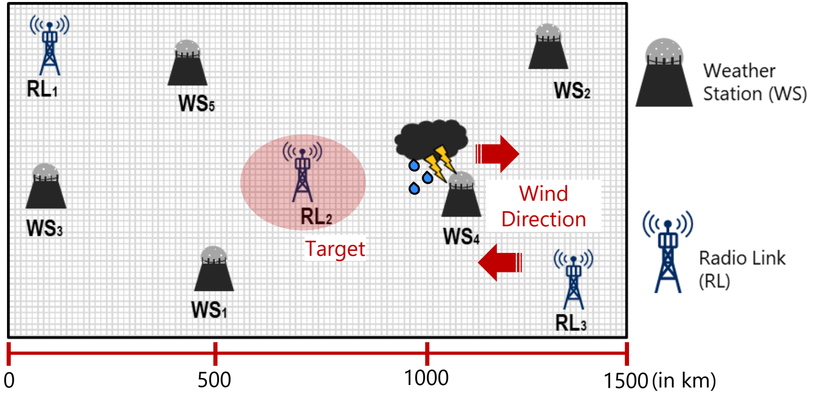
“Predict **Radio Link Failure** given **Weather Forecast Data**”

＜Raw Data Available＞

- Weather Forecast Data
- Radio Link Configuration
- Location Characteristics
- Relative Distances

Technical Challenges

Spatial and Temporal Correlations



Location of WS and RL are not the same!

Weather at the actual location of RL needs to be approximated based on weather forecast from different locations

Same forecast may have different impact at different times!

Actual weightage and whether at WS forecast is leading or lagging indicator depends on wind speed and direction

Technical Challenges

Spatial and Temporal Correlations

High Feature Space Dimensionality

Radio link (RL) failure doesn't depend only on weather but rather on combination of multiple types factors:

Weather Aspects

- Temperature
- Humidity
- Atmospheric pressure
- Rain/Snow precipitation etc.

Radio Link Configuration

- RL capacity
- Modulation scheme
- Link length
- Frequency band etc.

Topological Features

- Altitude of RL and WS
- Surrounding clutter type like urban, open-land, forest etc.

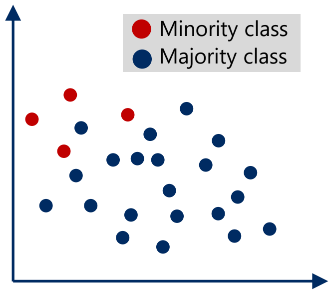
Results in **large feature space dimensionality** with a **mixture of categorical and numerical variables**

Technical Challenges

Spatial and Temporal Correlations

High Feature Space Dimensionality

Highly Imbalanced Dataset



In RL performance dataset, only **0.31%** of total samples resulted were RL failure

Famous/Standard Imbalanced datasets

Name of dataset	Minority Class %
Pima Indians Diabetes (Pima)	34.9%
Haberman Breast Cancer (Haberman)	26.5%
Thyroid Gland (Thyroid)	13.9%
Mammography	2.3%

Note: Evaluation metric is F1-score (and not accuracy)

Related Work

- Most existing research deals with adapting RL after detecting initial RL failure [1, 2]
- Following research work uses mainly 2 types of method for predictive analysis:

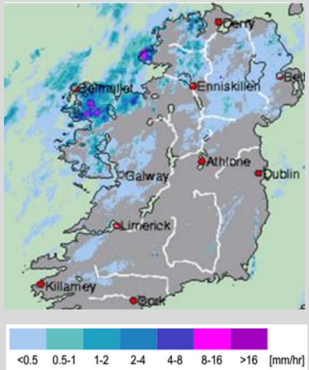
Type 1: Use of radar based rain echo maps [3, 4]

Step 1 — Acquire rain echo maps

Step 2 — Calculate signal attenuation based on analytical formulation

Step 3 — Predict future bit error rate and link failure probability

Example：Rain echo map



<MAJOR PROBLEM>

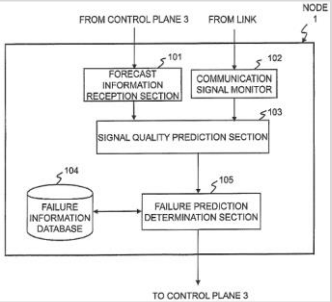
Low Prediction Capability as rain echo maps only available as "Now cast"

Type 2: Use forecast from nearest weather station
[US9246753B2, US2017/0026862A1]

Step 1 — Acquire weather forecast information

Step 2 — Prediction of radio signal quality using time series analysis

Step 3 — Raise warning if estimated signal quality is below threshold



<MAJOR PROBLEM>

Low Accuracy due to insufficient feature (weather, radio, topography) consideration

Data Pre-processing

80%

AI Model

(Random Forest)

20%

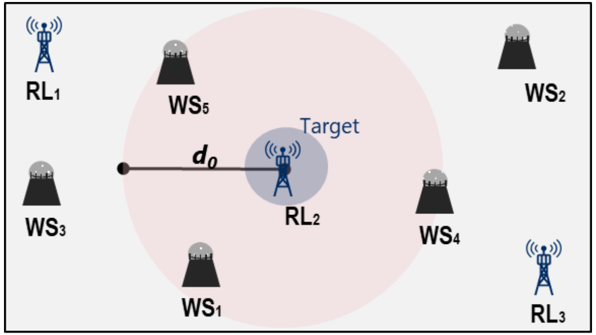
Selecting Weather Stations

- Use subset of WS which are 'd_0' distance apart from the target RL to predict RL failure

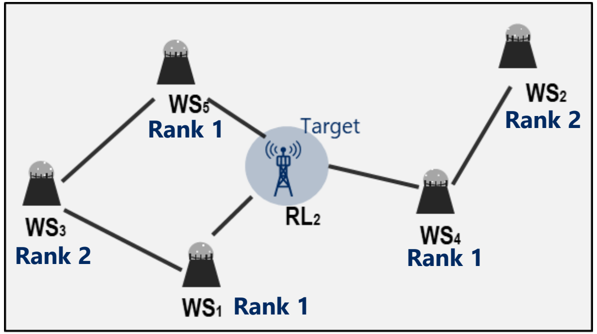
- Trade-off with value of 'd_0':

Very Large 'd_0' → **Low correlation** between forecast and radio performance

Very Small 'd_0' → **Less information captured** for RL failure prediction

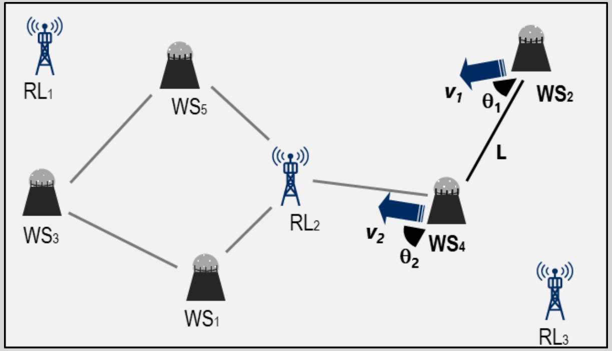


- Connect all nodes (all WS and target RL) with each other if distance between them is less than 'd_0'
- **Rank** of a WS is defined as the number of edges a WS is from target RL
- Choose the value of 'd_0' based on the:
  - Effective distances matrix
  - Resulting homological persistence



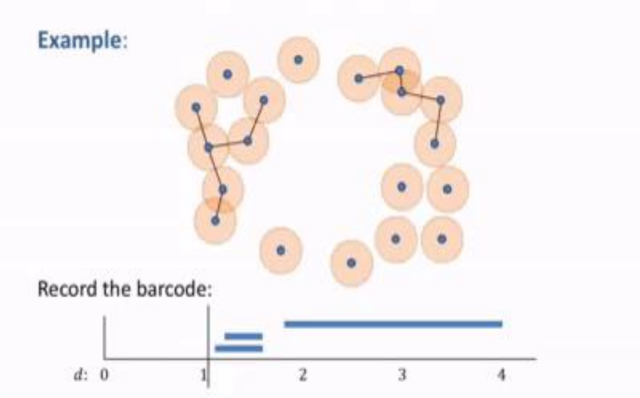
Selecting Weather Stations

Effective Distance



- Find the effective wind (say w_{eff}) along different nodes
- Then, effective distance is computed by scaling actual distance between WS and RL by $\frac{2}{1+e^{-k\times w_{eff}}}$
- Compute effective distance matrix

Homological Persistence

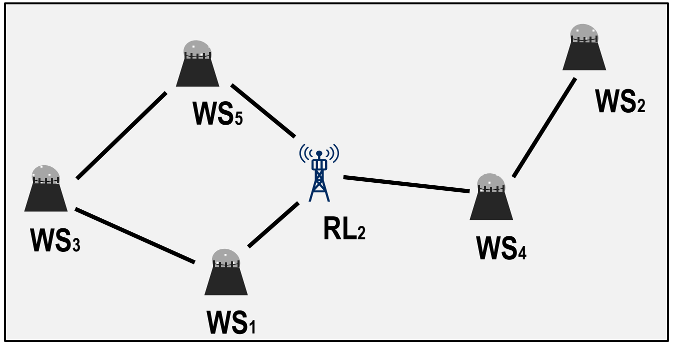


- As the value of *d*, nodes get connected to generate ring like features till entire graph is fully connected
- Identify the value of 'd_0' which captures most stable spatial feature using bar-code graph
- **Persistent homology** is a method for computing topological features of a space at different spatial resolutions.

Source: https://www.youtube.com/watch?v=2PSqWBIrn90

Data Integration

For a given day and given RL,



					RLF Label
Sample 1	RL_2 radio features	RL_2 Topography features	WS_1 forecast features	WS_1 Rank	TRUE
Sample 2	RL_2 radio features	RL_2 Topography features	WS_5 forecast features	WS_5 Rank	TRUE

⋮

and so on

⋮

We had data for 365 days and ~1000 RL, in total resulting ~3 million labelled samples

Feature decomposition and reduction

We have large number of both categorical and numerical features

Weather prediction	Clutter Class	Modulation scheme
SCATTERED CLOUDS	LOW-DENSE URBAN	C-QPSK
HEAVY THUNDERSTORM	OPEN IN URBAN	512QAM(QO)
SLEET	OPEN LAND	128QAM
HEAVY RAIN	AVERAGE-DENSE URBAN	256QAM(Q)
OVERCAST CLOUDS	LOW-DENSE URBAN	1024QAM
LIGHT SNOW	INDUSTRIAL & COMMERCIAL	256QAM
FEW CLOUDS	SPARSE TREE	1024 QAM
MISTY	BUILTUP-VILLAGE	4QAM
FOGGY	INDUSTRIAL & COMMERCIAL	128QAM-XPIC
LIGHT RAIN SHOWERS	AVERAGE-MEDIUM URBAN	128QAM(Q)
WINDY	AVERAGE-SPARSE URBAN	32QAM
HOT DAY	LOW-MEDIUM URBAN	2048QAM
CLEAR SKY	GREEN HOUSE	256QAM(QO)
⋮	⋮	⋮

and many other categorical features

One-hot encoding results into curse of dimensionality

Feature decomposition and reduction

We have large number of both categorical and numerical features

To effectively tackle this problem:

I. Binary decomposition

- One-hot encoding unsuitable due to resulting feature vector's extremely large dimensionality
- Instead we decompose categorical features into multiple meaningful binary variables

Clutter Class

	Height	**Density**	**Man made**
LOW-DENSE URBAN	**0**	**1**	**1**
OPEN LAND	0	0	0
SPARSE TREE	0	0	0
BUILTUP-VILLAGE	0	0	1
LOW-MEDIUM URBAN	0	0	1
GREEN HOUSE	0	0	0

Similar approach was used for categorical variables as well

Feature decomposition and reduction

We have large number of both categorical and numerical features

To effectively tackle this problem:

I. Binary decomposition

- One-hot encoding unsuitable due to resulting feature vector's extremely large dimensionality
- Instead we decompose categorical features into multiple meaningful binary variables

II. Feature Reduction

- To further reduce dimensionality of input feature space, we rank features based on their importance
- Iteratively remove least important feature till we observed drop in F1 score

Example:

```
Variable: eff_distance Importance: 0.11
Variable: temp_max_day1 Importance: 0.08
Variable: humidity_max_day1 Importance: 0.07
Variable: distance Importance: 0.07
Variable: speed Importance: 0.07
Variable: tip_FAR Importance: 0.06
Variable: clutter_type Importance: 0.04
Variable: modulation_intensity Importance: 0.04
Variable: type_NEC Importance: 0.04
Variable: density Importance: 0.03
Variable: rain Importance: 0.03
Variable: cloud Importance: 0.03
Variable: intensity Importance: 0.03
Variable: adaptive_modulation_Enable Importance: 0.02
```

Least Important

Iteratively remove least important feature till we observe drop in F1 score

Dataset Re-balancing

Original dataset was highly imbalanced with ratio of **Minority : Majority :: 1 : 320**

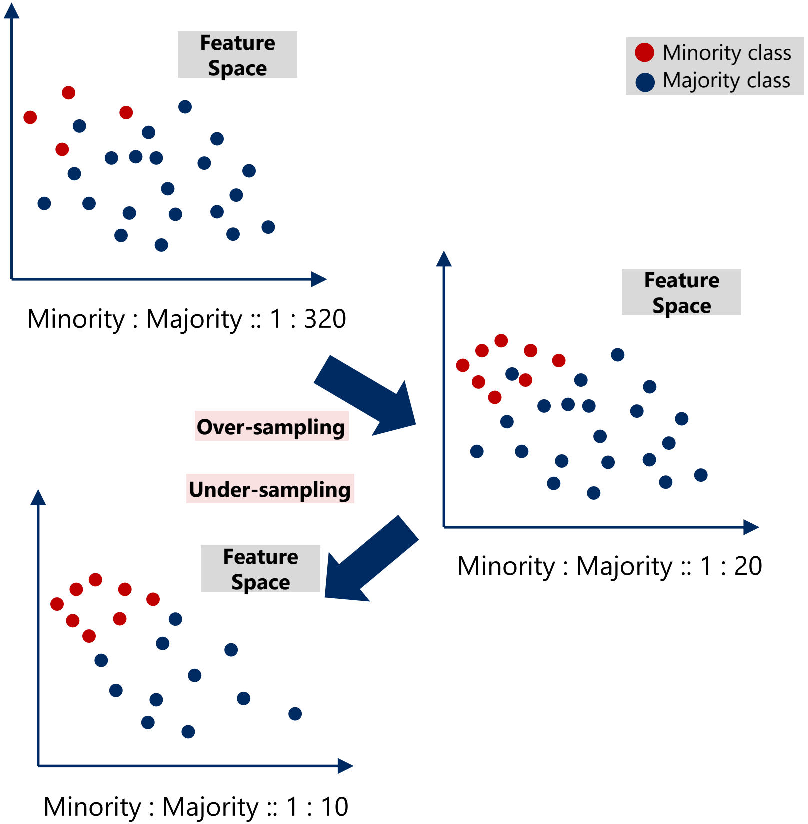
Oversampling

- Oversample minority class samples by generating synthetic samples using SMOTE (Synthetic Minority Over-sampling Technique) [5]

Under sampling

- Under-sample majority class samples randomly

Final dataset was rebalanced to get ratio of **Minority : Majority :: 1 : 10**



AI Model and Inference

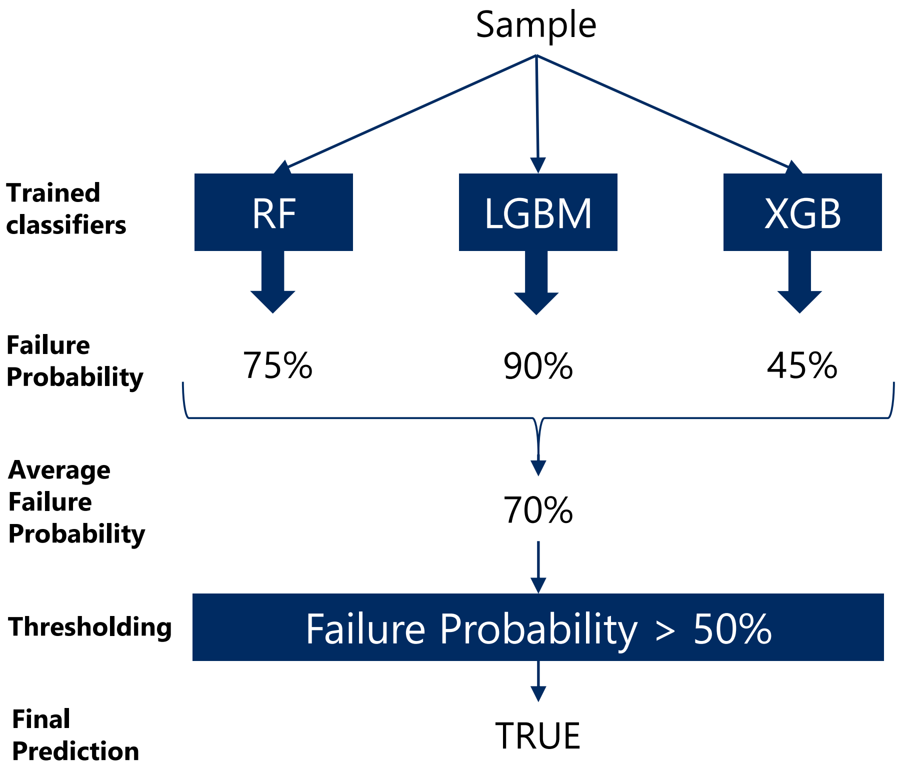
Particularly we used a combination of

- Random Forest
- LGBM (Light Gradient Boosting Machine)
- XGB (Extreme Gradient Boosting Machine)

Ensemble method decreases variance, bias, and improves prediction!

Decision Tree based classifiers were chosen as:

- Good performance with high dimensional data
- Feature importance
- Fast training with low computational requirements
- Prediction confidence level
- Explainable AI (combined with SHAP analysis)



Summary

Given the real-world data & problem complexity, following challenges were tackled:

Problems	Solution Employed
I. Spatial and temporal correlations Not all the WS forecast effect the RLs in the same way. The weather conditions such as wind direction and topology creates also create a difference	**Graph based WS ranking** Graph Structure using homological persistence and effective distances to choose WS for RLF prediction
II. High feature space dimensionality Large number of features related to RLs, weather forecasts, topography makes it difficult task to predict RL failure reliably	**Feature decomposition and reduction** Feature decomposition and Iterative feature reduction based on the importance level of feature for AI model
III. Highly imbalanced dataset RL failures are very rare with only <0.3% of data samples having RL failure (=True). This makes it very difficult for AI model to predict the failures reliably	**Dataset rebalancing** Generate synthetic data resembling minority class and under sampling majority class to rebalancing the dataset

Evaluation

- Divided dataset into 70 : 30 :: Training : Test datasets.
  Evaluation on following metrics:

$$Precision = \frac{TP}{TP + FP} \qquad Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Recall = \frac{TP}{TP + FN} \qquad F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

	Predicted RLF	Predicted No RLF
True RLF	**TP**	**FP**
True No RLF	**FN**	**TN**

Confusion matrix

- Evaluation for 1 day ahead prediction:

	Accuracy	Precision	Recall	F1 score
Proposed Method Performance	**1.00**	0.72	**0.86**	**0.78**
Without effective distance	**1.00**	0.59	0.85	0.69
Without feature decomposition (using one-hot encoding)	**1.00**	0.00	0.00	0.00
Without re-balancing dataset	**1.00**	**0.80**	0.77	**0.78**

- **Results in lower precision i.e., more FP**
- **Large feature space results in no learning**
- **Better precision at the cost of lower recall**

Concluding Remarks

We presented a method to predict RL failure using weather forecast data.

Some of the **salient features** are:

- Fast and low computational complexity training
- Prediction confidence value
- Explainability of predictions

Uses of the accurate RL failure prediction **(Inputs from the host – 'Turkcell')**

- Prediction can be used by self-organizing networks (SON) to change operational frequency, tilt etc.
- Accurate predictions can save man-hours as many RL changes require field engineer in person
- Understand what kind of failure results from what kind of weather conditions and thereby,
  - Help in field deployment and planning in future in long-term
  - Prepare in advance for seasonal changes

References

1. N. Javed, E. Lyons, M. Zink and T. Wolf, "Adaptive Wireless Mesh Networks: Surviving Weather without Sensing It," 2013 22nd International Conference on Computer Communication and Networks (ICCCN), Nassau, 2013, pp. 1-7, doi: 10.1109/ICCCN.2013.6614102.
2. Abdul Jabbar, Justin P. Rohrer, Victor S. Frost, James P.G. Sterbenz, “Survivable millimeter-wave mesh networks”, Computer Communications, Volume 34, Issue 16, 2011, Pages 1942-1955, ISSN 0140-3664,
3. A. Jabbar, J. P. Rohrer, A. Oberthaler, E. K. Cetinkaya, V. Frost and J. P. G. Sterbenz, "Performance Comparison of Weather Disruption-Tolerant Cross-Layer Routing Algorithms," IEEE INFOCOM 2009, Rio de Janeiro, 2009, pp. 1143-1151, doi: 10.1109/INFCOM.2009.5062027.
4. Rak, J. A new approach to design of weather disruption-tolerant wireless mesh networks. Telecommun Syst 61, 311–323 (2016). https://doi.org/10.1007/s11235-015-0003-z
5. H. M. Nguyen, E. W. Cooper, K. Kamei, “Borderline over-sampling for imbalanced data classification,” International Journal of Knowledge Engineering and Soft Data Paradigms, 3(1), pp.4-21, 2009.

\Orchestrating a brighter world

NEC

Questions?

Comments?

Suggestions?