A Prediction-based Online Cost Optimization Algorithm for 5G Vertical

Jyoti Tiwari, Shashwat Kumar and Antony Franklin A.
Indian Institute of Technology Hyderabad, India

IEEE International Conference on Advanced Networks and Telecommunications Systems

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Introduction

- Vertical industry focuses on a **particular** niche, eg : automotive, education, manufacturing and real-estate.
- Fast-paced change in business ecosystems move towards digitalization of Vertical Industries.
- **5G** will be a **major** technology in **growing** industrial digitalization, such as immersive gaming, autonomous driving, remote robotic surgery and augmented reality.

Requirements of Vertical and it's Sharing

Each verticals have **different service requirements** in terms of Key Performance Indicators (**KPIs**) such as throughput, delay and reliability, and network requirements such as isolation and special authentication.

- Through SLA, requirements are shared between the operator and the Vertical, which determines the **guaranteed** level of performance and corresponding **cost**.

5G Enabler for Verticals and Need of Cost Optimization

- Mechanism of **network slicing** is one of the key **enablers for 5G** networks to support the architecture.
- 5G network slicing is a network **architecture** that **enables** the multiple independent **logical networks** on the same physical network infrastructure to **provide** telecommunication services.
- With **increasing demand** of verticals, resource requirement increases.
- Due to limited resources and higher cost, effective allocation of resources are required.

Main Obstacle in Cost Optimization

Dynamic traffic demand of verticals.

Online Problem

Based on the concept of **online algorithm**, which processes its input piece-by-piece **without having the entire input** available from the start.

The goal of this problem is to **minimize/maximize the objective** function without having the entire information available at start.

Vertical's Resource Reservation

Verticals can reserve resources in two ways :

- Long-time reservation of resources (SLA update/ Long-term Plan)
- On-demand reservation of resources (Short-term Plan)

Per-day charge of resources **decreases** as the leasing period **increases** due to discounted long-term plans.

Pricing of Bandwidth Resources by Multiple Service Providers

Service Provider	Plan (p)	Contract Period (w)	Price (s in $)	Per day rate (r in $)
Jio	2 GB/day	28 days	3.4	0.12
Jio	2 GB/day	84 days	8.2	0.11
Airtel	2 GB/day	28 days	4.09	0.14
Airtel	2 GB/day	84 days	9.6	0.11

Motivation

- Traffic variation of Verticals in 5G is **unknown** to the Verticals and Provider.
- Rate of resources decreases with increase in leasing period, how we can utilize this information for cost optimization in dynamic traffic environment.

Problem Formulation

Select a cost-effective plan to support the uncertain traffic demand of verticals via mapping with a multi-slope variant of ski-rental problem.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Online Ski-Rental Problem

- A skier needs to decide between **buying skis at cost b** and **renting** them at the cost of 1 per day.
- But, the skier **does not know** the length of the ski season in advance.
- **Deterministic** strategy for the skier is to rent for b-1 days and buy at b^{th} day, which achieves best worst case **competitive ratio of approx 2.**
- **Randomized** strategy achieves best worst case **competitive ratio of approx 1.58 .**

Proposed Prediction-based Online Cost Optimization Algorithm for Verticals

- Data Preprocessing Phase
- Traffic Prediction Phase
- Plan Selection Phase

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Data Preprocessing Phase

- **Telecom Italia** has open-sourced a user interaction dataset from the city of **Milan and the Province of Trentino**, which contains **CDRs for each 10 minutes** .
- Call Detail Records (CDRs) are collected from each grid of Milan and Trentino.
- Each time a user initiates a telecommunication interaction, a Radio Base Station (RBS) starts a new CDR recording.
- Mapped the each grid’s CDRs traffic to the traffic of network slice.
- Generated a **normalized aggregated per day** traffic dataset using CDRs of each day.

Dataset

```
In [16]: data
```

Out[16]:

	1	2	3	4	5	6	7	8	9	10	...	91	92
0	93.724525	86.762950	114.880818	267.477315	278.717637	384.986951	314.069895	139.149709	113.774670	163.453650	...	207.879692	143.800920
1	243.589731	234.137541	261.514191	546.668701	557.610767	763.080380	681.351742	311.303674	256.413114	315.973322	...	463.630456	363.593934
2	112.172493	109.916621	124.632083	301.576681	324.152691	439.332003	388.829577	192.156226	175.431736	201.202183	...	243.587628	273.646687
3	158.303830	151.822996	160.461245	317.717487	354.137755	479.632251	431.861673	208.858100	249.719621	271.944320	...	295.400164	338.311128
4	93.780078	97.841884	115.968455	192.387388	212.470888	286.552975	253.793839	142.697457	143.608276	166.551650	...	215.348924	197.136609
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1292	84.224500	79.730052	88.224452	185.832099	201.310738	262.708440	245.796943	135.936380	95.720888	113.456286	...	343.385628	260.082297
1293	123.710273	122.146803	113.758355	256.952113	280.840076	345.634485	314.588165	192.517047	117.309073	159.525139	...	421.074217	350.696610
1294	93.986241	96.319440	90.547909	174.411435	191.030156	241.004327	212.815626	115.359299	77.806505	110.420934	...	310.018248	243.040804
1295	114.859605	149.535939	190.271113	232.030345	235.671418	324.936449	299.596130	160.559475	106.293249	157.338722	...	400.829646	315.012177
1296	29.630386	39.928679	57.652265	62.202624	60.007671	86.672490	82.997526	53.269741	35.897018	48.652206	...	104.631513	91.869955

1297 rows × 100 columns

Traffic Prediction Phase

- **LSTM** model is used to **predict future traffic** to improve the effectiveness of the decision-making.
- LSTM is a special kind of RNN and is widely used for time-series prediction due to its ability to learn long-term context.
- The accuracy of the LSTM model improved by retraining continuously on recently available data.
- **Continual learning** is beneficial in an environment where the data trend keeps changing.

LSTM Model Configuration Parameter

Activation Function	ReLU
Optimizer	Adam
Number of training Dataset	900 days
No. of Layers	256
Epochs	50
Time Steps	7 days

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Plan Selection Phase

- Multi-slope online plan-selection algorithm utilizes prediction information of LSTM model to chooses a suitable plan from a set of plans.
- Algorithm decides whether to on-demand a plan or update the existing plan..
- Computes the online cost and competitive ratio for competitive analysis.
- With new stream of data, reiterate all phases to improve the accuracy of model and better optimize the cost.

Competitive Analysis

- **Competitive analysis** is a method invented for analysing performance of online algorithm.
- The performance of an online algorithm is compared to the performance of an optimal offline algorithm that can view the sequence of requests in advance.
- An algorithm is competitive if its competitive ratio is bounded.

Proposed Algorithm

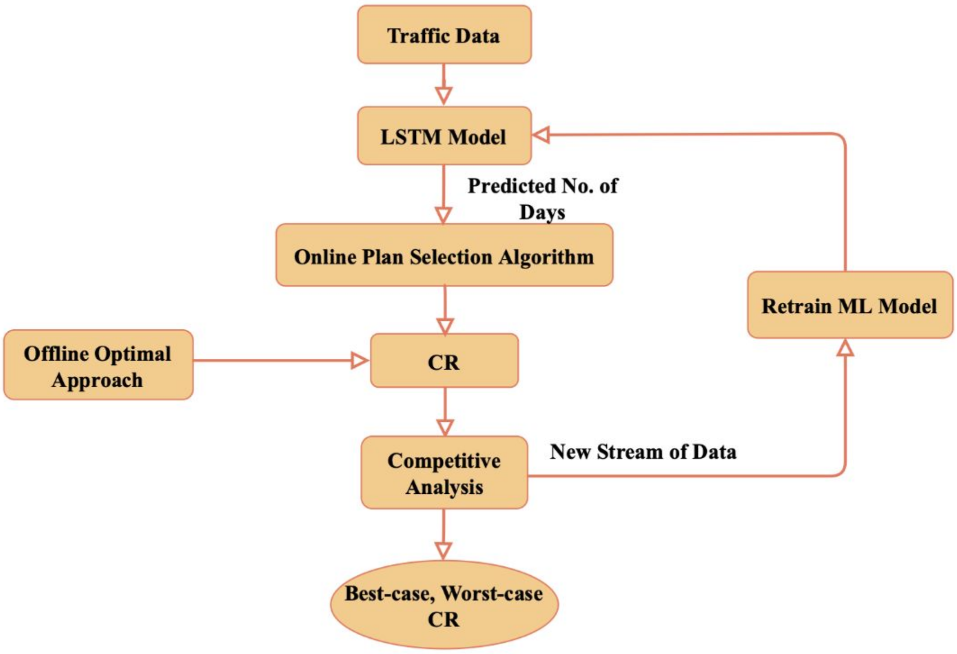
Algorithm 1: Online plan selection algorithm

Input: $LSTM_{prediction} \rightarrow \{B_p, m\}$, $Offline_{optimal} \rightarrow \{C_{opt}\}$, B_c, C_c.
Output: C_o, CR
while *(w days stream of data)* **do** // competitive analysis
  if $B_p > B_c$ **then** // comparison between predicted and acquired bandwidth
    if $m * r_i >= s_i$ **then** // comparison between on-demand cost and SLA update cost
      reserve s_i resource at start;
      $C_o = s_i$;
    else
      on-demand additional resource at r_i;
      $C_o = C_c + m * r_i$;
    end
  else
    $C_o = C_c$;
  end
  $CR = \frac{C_o}{C_{opt}}$
end

Simulation Parameters

On-demand duration	1 Day
SLA Contract Period (w)	28 Days
Current Bandwidth B_c	160 GB
Current Cost C_c	\$932

Flow of Proposed Model



Offline Optimization Algorithm for Cost Computation

$$\underset{f}{\text{minimize}} \quad f(x, y) = \sum_{t=0}^{T} \sum_{i=1}^{n} (x_{ti} r_i + y_{ti} s_i) \tag{1a}$$

$$\text{subject to} \quad \sum_{i=1}^{n} p_i * x_{ti} + p_i * y_{ti} >= d_t \forall t, \tag{1b}$$

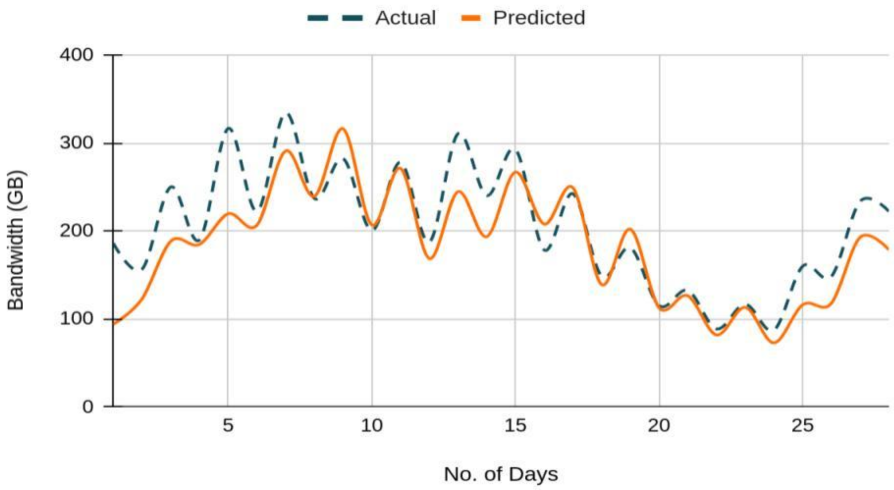
$$x_{ti}, y_{ti} \in \{0, 1\}, \tag{1c}$$

$$d_t, p_i, r_i, s_i \geq 0. \tag{1d}$$

Where,

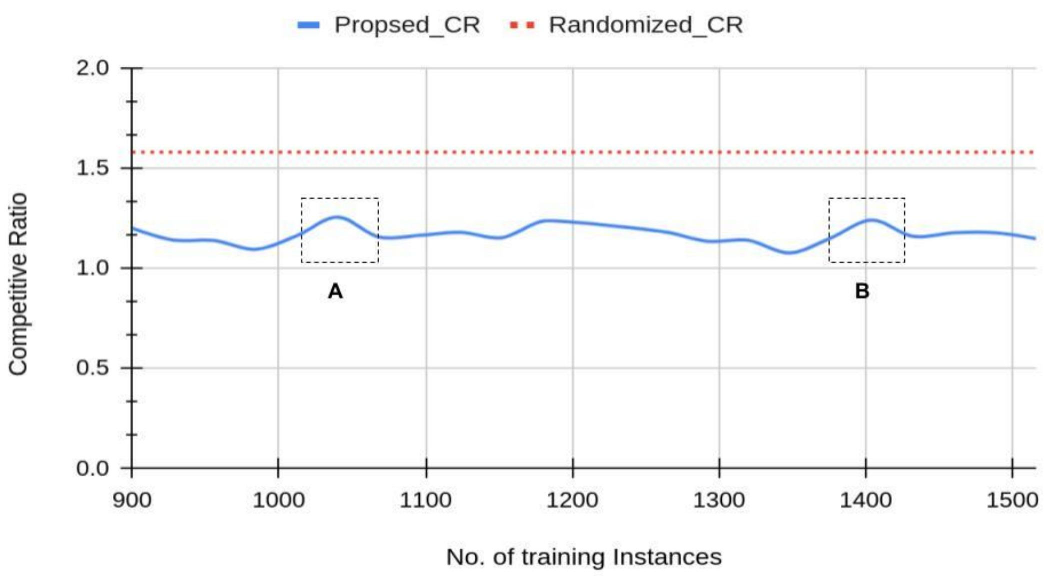
- $\mathbf{x}_{ti}$ and $\mathbf{y}_{ti}$ are the binary decision to indicate the selection of short-term or long-term plan
- On-demand cost $\mathbf{r}_i$
- Long-term plan cost $\mathbf{s}_i$
- bandwidth demand $\mathbf{d}_t$ enters the system at time t.
- $\mathbf{p}_i$ corresponds to the allocated bandwidth.

Prediction on Dataset diagram



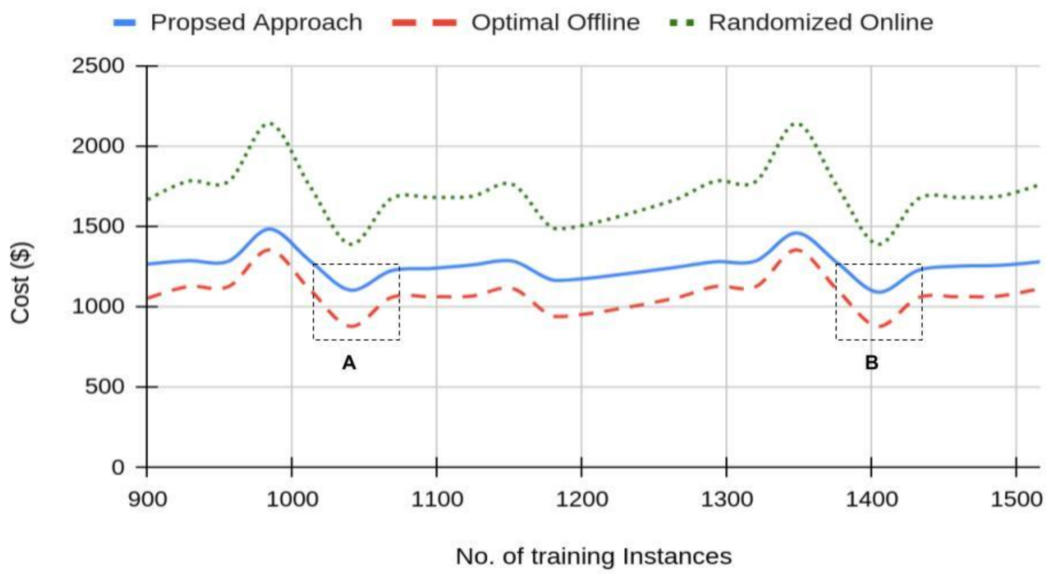
X-axis represents to the number of days.
Y-axis represents to the bandwidth in GB.

Competitive Ratio Analysis



- Competitive analysis of proposed algorithm with best worst-case CR of randomized algorithm.
- The **dynamic change** in **demand** and **low prediction** accuracy caused the worst case CR at point A and B.

Cost Comparison Analysis



Summary

- This work addresses the plan selection problem for vertical industries with dynamic traffic demands.
- The proposed algorithm can help the different industry verticals to optimize the operational cost by choosing a network slice plan.
- The performance evaluation on the real-world dataset suggests that the proposed algorithm improves over **randomized algorithm by 20%** and **deterministic algorithm by 37%** in worst case scenario.

Acknowledgement

This work was supported by the Department of Telecommunications, Ministry of Communications, India as part of the “Indigenous 5G Test Bed” project.


THANK YOU