AI accelerates readiness for 5G data billing at China Unicom

Who?

Since the launch of 4G, operators the world over have seen explosive growth in traffic as data limits have risen, prices have fallen, and the networks’ capacity and speed increased. This was China Unicom’s experience, accompanied by soaring number of messages concerning credit control requests (CCRs), generated by the Diameter Credit-Control application protocol which defines a charging mechanism for pre-paid users. To do this, it uses a credit-limit control to implement session- and event-based charging. Likewise, the rise in traffic also produces huge numbers of call detail records (CDRs).

What?

Needed to reduce the resources required for billing and find a future-proof solution for billing and its associated systems ready for the bigger data surge from 5G

How?

Use of cloud-native, artificial intelligence platform to predict usage, reducing the number of CCR messages by 63% and CDRs by 52%

Results

Savings of 50% on bandwidth and computing resources, as well as more than halving the amount of storage required
There is an urgent need to reduce the strain on and soaring costs of storage, bandwidth and systems associated with billing. At the same time, operators need to improve operational efficiency as well as to implement global governance strategies and invest heavily in new hardware and especially software technologies – and all with without negatively affecting users’ experience.

Against this backdrop, China Unicom decided to explore how artificial intelligence (AI) could be used to address these challenges. In parallel it wanted to investigate how edge computing could reduce the pressure on the network elements across the infrastructure and their associated systems, as well as help with the expected burst of traffic at the edge.

### Taking the strain

In 2019, China Unicom set up a project – involving a customer base of 5.3 million – with Si-Tech, a telecom solution provider in China. Whereas conventional approaches use a fixed data quota policy for each user, the partners developed on AI-powered application on a containerized – cloud-native – platform. The platform has unified internal packaging and standardized external interfaces, for rapid deployment of applications.

The AI technology was in the form of a neural network, which was used to build intelligent quota models based on static user tags, which contain customer’s personal details, and user behavior tags, which track customers’ interactions and preferences.

The companies used two of TM Forum’s Open APIs – for alarm and quota management – to simplify and speed up the implementation of the platform and intelligent quota model through standardized integration, end to end. In this way, the operator was quickly able to feed huge amounts of data into the modeling application. This included original data source and billing system data from China Unicom’s big data platform, which is a unified management system for data.

As a first step, the operator used big data analysis to determine the most effective way to train the AI. The AI monitors how the models run, then ‘learns’ from outcomes, and through continuous iteration of the models optimizes their performance.

From the tags, the models predict data about users, services and time periods, and assigns an intelligent data quota, which can be matched against actual usage in real-time. This predictive capability reduced CCR-related messages by 63% in Shandong province, which equates to 964 million fewer CCR messages every day (also see results below).

By the end of the trial, China Unicom had cut, by 50%, the bandwidth and computing resources needed to handle the billing and CDRs efficiently. This massively reduced the amount of processing and response delays from network elements, the online charging systems and the business service inquiry system, such as for balance and account inquiries, which is related to 5G billing, which were not part of the trial. All of which also meant that customers noticed that the service had improved.

Other systems that are connected to the 5G billing system, like big data platforms, will also save a lot of storage space and improve the overall efficiency of the systems’ processing.

### Using AI to build models
While the data analytics are used to mine the data held in core billing systems, and to identify that value through business operations, it is the AI which realizes that value. The AI-powered intelligent billing in this project makes a strong use case for operators to employ AI technology to empower their cloud-based systems.

This project also covered edge computing. By dynamically allocating traffic quotas based on users’ online behaviour, the amount of data is reduced significantly. This means that after data processing at the edge, only the useful data is sent to the billing system.

In future, new use cases will be explored such as: automatic switching of online and offline charging based on the user’s credit, balance and other information; and intelligent control based on the user’s 5G context, such as low-latency scenarios. China Unicom will also consider putting part of its computing capabilities on the edge of the network.
The project showed that China Unicom can reduce the number of billing requests and increase processing performance correspondingly, so that the same systems can support more users and their requests.

- Daily CCR messages reduced by 964 million equating to a 63% fall from 1.57 billion, which reduced the monthly total of CCRs to 28.917 billion (30 days x 796 million).
- The daily number of CDRs fell by 52%, to 468 million from 900 million, which reduced the monthly total to 14.040 billion (30 days x 468 million)
- Average amount of storage needed daily before going online was 3.86TB, which fell by 1.84TB when operations moved to the containerized platform
- This reduced monthly storage by 55.2TB (30 days x 1.84TB)
- The system stores 13 months’ of data, hence over that period the amount of storage required has fallen by 717.7TB (13 months x 55.2TB)

As China is the world’s third largest country in terms of landmass and has the largest population, the success of this project suggests if this platform and approach can solve the challenge in China, then they can be implemented anywhere on the planet by operators to address the same issues.