Demo: Immersive Network Operations using 5G-Enabled XR Headsets and MCP

Abhiram Elango
Northeastern University
Boston, MA, USA

Ankita Mandal
Northeastern University
Boston, MA, USA

Eduardo Baena
Northeastern University
Boston, MA, USA

Dimitrios Koutsonikolas
Northeastern University
Boston, MA, USA

Abstract

We demonstrate a fully voice-driven troubleshooting assistant for 5G edge networks using XR headsets with integrated 5G connectivity. The human operator describes symptoms and navigates diagnostics entirely through speech; the LLM responds via synthesized voice while simultaneously rendering relevant metrics as AR overlays. This multimodal feedback, audio explanation paired with visual data, enables hands-free, eyes-free operation in isolated field environments where traditional interfaces are impractical

ACM Reference Format:
Abhiram Elango, Ankita Mandal, Eduardo Baena, and Dimitrios Koutsonikolas. 2026. Demo: Immersive Network Operations using 5G-Enabled XR Headsets and MCP. In *The 27th International Workshop on Mobile Computing Systems and Applications (HotMobile '26), February 25–26, 2026, Atlanta, GA, USA.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3789514.3796260

1 Introduction

Field troubleshooting of 5G infrastructure demands operator mobility and situational awareness incompatible with laptop-based tools. Technicians inspecting equipment cannot simultaneously read terminal output; environments lacking cloud connectivity preclude standard LLM interfaces. We present a voice-first diagnostic assistant where interaction occurs through natural speech, with the LLM providing spoken explanations synchronized with visual metric displays.

The system extends our prior MCP-based architecture [1] with: (1) edge-local speech processing (STT/TTS) for bidirectional voice interaction, (2) AR overlays rendering metrics while the operator listens, and (3) physical button confirmation ensuring deliberate human approval for each diagnostic action.

HotMobile '26, Atlanta, GA, USA

ACM ISBN 979-8-4007-2471-8/26/02
https://doi.org/10.1145/3789514.3796260

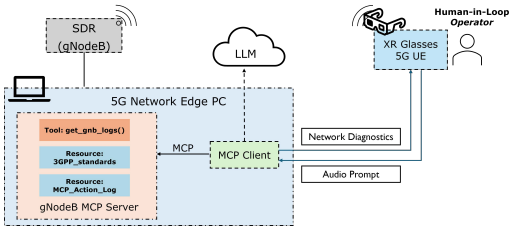

Figure 1: System Architecture

2 System Architecture

Figure 1 shows the architecture. The XR headset captures voice via microphone; the edge node transcribes (Whisper), reasons (LLM), and synthesizes responses (TTS). Metric visualizations render on the AR display synchronized with spoken explanations. All processing executes locally without external connectivity.

3 Interaction Model

The operator converses naturally with the LLM, which proposes diagnostic steps. When the LLM suggests querying a specific log or metric, a confirmation prompt appears on the AR display. The operator must press a physical button on the headset to approve, ensuring the action reflects deliberate human decision rather than conversational ambiguity.

Example: Operator reports "Video calls are freezing." The LLM verbally suggests checking RLC retransmissions; AR displays `[Query RLC?]` with a button prompt. Upon physical confirmation, metrics render as gauges and time-series while the LLM narrates findings.

Visual feedback: Status indicators (green/yellow/red), metric gauges with thresholds, and time-series plots appear alongside spoken explanations, reducing cognitive load versus reading logs.

4 Demonstration Setup

Hardware: OpenAirInterface gNodeB (B210 SDR), 5G-enabled XR headset, edge laptop, internet access to cloud-hosted LLM (Claude Sonet 4.5).

References

[1] Eduardo Baena, Ankita Mandal, and Dimitrios Koutsonikolas. [n. d.]. Demo: Human-in-the-Loop Agentic Reconfiguration of Edge 5G Networks via Dual-MCP and LLM Reasoning.