LLM for 5G: Network Management

Ali Mamaghani*, Ali Nourian*, Negin Mohtaram*, Alireza Shokrani*, Mohsen Nasiri*,
Sepehr kazemi*, Alireza Mohammadi†, Navid Nikaein† and Babak Hossein Khalaj*
*Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran, Email: {ali.mamaghani, ali.nourian, neg.mohtaram, alireza.shokrani, seyedmohsen.nasiri21, sepehr.kazemi, khalaj}@sharif.edu
†EURECOM, Sophia-Antipolis, France, Email: {alireza.mohammadi, navid.nikaein}@eurecom.edu

***Abstract*—This demonstration explores the effective incorporation of Large Language Models (LLMs) in 5G network automation, significantly enhancing the process of private network creation without the need for extensive networking or programming knowledge. The demonstration comprises a user-friendly chatbot, adept at translating everyday English queries into actionable 5G commands, and LLMs serving as generative AIs to dynamically generate configurations tailored to the 5G network environment. Our approach relies on the practical application of text classification models and an adaptive feedback loop mechanism, allowing the system to recognize patterns, refine commands iteratively based on real-world execution errors, and strengthen overall resilience. The demonstration comprises two interlinked parts: the development of a chatbot accessible to a wide audience and the utilization of LLMs for dynamic configuration generation. The goal is to exemplify the accessibility of 5G network creation to a broader user base, eliminating barriers for those without extensive technical backgrounds. This comprehensive approach not only simplifies the process but also enhances the system's responsiveness to the varied challenges presented by real-world 5G network scenarios, laying the groundwork for a future where 5G networks are both accessible and adaptable across diverse user profiles.**

I. Introduction

In the transition from 4G to 5G and beyond, the concept of private mobile networks attracted the community's attention. However, creating private networks needs knowledge of networking systems and programming. On the other hand, recent advancements in the field of Artificial intelligence (AI), especially the emergence of LLMs, have impacted the whole aspects of communication systems. In this demo, we aim to show how we can leverage LLMs to create and interpret a 5G network. As a matter of fact, people without knowledge of networking systems and programming can also have their own 5G private network. The demo contains two parts: 1- Developing a chatbot which is capable of translating everyday English queries to be understandable for our platform. 2- Using LLMs as a generative AI so as to create configurations for our platform. In the former, we fine-tune some models which are able to distinguish the category of request including but not limited to install, test, and login, and generate the command appropriate to the platform. In the latter, we finetune models to generate configuration files and use system error feedback to get the best result.

Language is a cornerstone of human communication, and as artificial intelligence seeks to engage in coherent communication, LLMs have emerged as pivotal tools for text processing and generation. The surge in demand for machines to handle complex language tasks, including translation, summarization, information reval, and conversational interactions, has fueled the development of LLMs. Pre-trained language models (PLMs), a category within LLMs, have exhibited remarkable generalization abilities for text understanding and generation tasks when trained in a self-supervised setting on extensive text corpora. Notably, the performance of PLMs further improves through fine-tuning for downstream tasks, surpassing models trained from scratch. The evolution of LLMs is evident in the increasing number of released models. Notably, models like GPT-3.5 showcased impressive performance.

Moreover, the success of LLMs has ignited research into fine-tuning methods, offering a cost-effective and accessible means to enhance model performance. These methods enable more control over LLM outputs, facilitating tailored adjustments that significantly improve the model's capabilities and responsiveness to specific tasks and domains.

In our work, we introduce significant contributions to the field, including:

- **Novel Use of LLMs:** We showcase the novel application of LLMs in 5G network creation, demonstrating their pivotal role in making the process accessible to individuals without networking or programming expertise.
- **Innovative Integration of Fine-Tuned Models:** Our methodology incorporates fine-tuned models for text classification, enabling the accurate categorization of user-provided descriptions into specific 5G network commands. This approach distinguishes our work from traditional methods, allowing for a more streamlined and efficient process.
- **Adaptive Feedback Loop for Real-World Resilience:** The introduction of an adaptive feedback loop in our methodology represents a breakthrough. This feature refines generated commands iteratively based on real-world execution errors, enhancing the system's resilience and adaptability to dynamic 5G network scenarios.

These contributions collectively highlight our efforts to revolutionize 5G network creation, making it more accessible, efficient, and responsive to diverse real-world situations.

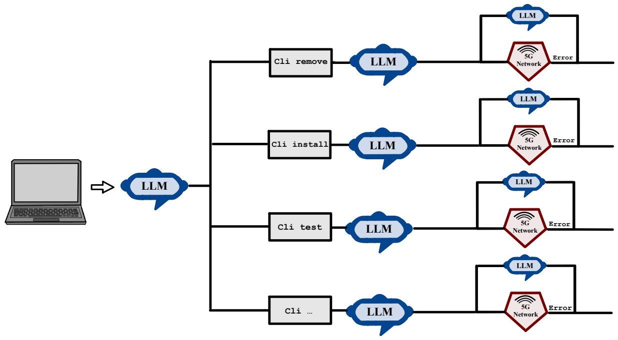

Fig. 1. The diagram illustrates the progression of the user's request within our system in a two-step process. Initially, the request undergoes classification by an LLM model fine-tuned for this purpose. Subsequently, the LLM model corresponding to the identified category generates platform-specific commands or configuration files. Finally, a feedback loop directs potential system errors to a dedicated model for error correction.

II. Implementation

Our approach leverages LLMs in the context of 5G network automation. The overall system architecture is designed to seamlessly translate user-provided textual descriptions into executable commands and generate relevant configuration files for 5G networks. This methodology involves a multi-stage process, as illustrated in Figure 1.

The first stage of our methodology involves a text classification model. LLMs can be effectively employed for text classification due to their adeptness in leveraging statistical relationships between words and phrases, allowing them to comprehend the intricate contextual nuances of language. This proficiency enables them to discern patterns indicative of different categories. Our fine-tuned model is trained to categorize user-provided descriptions into distinct classes, each representing a specific type of command within the 5G network environment. These categories serve as the branching points for subsequent processing. Based on the output of the text classification model, the system bifurcates into multiple branches, each corresponding to a recognized category.

Each branch is associated with a category-specific language model, fine-tuned on a specified dataset related to the recognized category. In this phase, the focus is on generating executable commands and configurations specific to the 5G network. The result from each specialized language model falls into one of two categories: a command crafted in a language understandable by 5G networks or a configuration file in which our 5G network's parameters are set in based on the user's request. This leads to the next step where LLMs demonstrate their proficiency. LLMs can be employed adeptly in prompt generation by crafting text consistent with a given prompt, enabling them to generate prompts for other LLMs to perform specific tasks. The key lies in the LLMs' capacity to produce clear, concise prompts containing all necessary information for the designated task. This ensures that the generated commands align with the syntax and semantics expected by the 5G infrastructure.

The generated commands and configuration files are transmitted to the 5G network for execution. In the event of an error during execution, an adaptive feedback mechanism comes into play. The error information, along with the initially generated command or configuration file, is used to iteratively refine the command for subsequent attempts. This dynamic process continues until a successfully executable command or configuration file is generated.

By integrating a feedback loop, our methodology adapts to real-world scenarios, continuously learning from errors encountered during execution. This iterative refinement process enhances the system's resilience and ability to respond effectively to a wide range of potential errors.

III. Conclusion

In conclusion, our research introduces an approach that makes 5G network creation more user-friendly, utilizing LLMs. By employing LLMs, we enhance the accessibility of the process, breaking down barriers for individuals without specialized technical backgrounds. The implementation of fine-tuned models for text classification ensures precision, simplifying the generation of 5G commands.

The incorporation of an adaptive feedback loop enhances the system's resilience in real-world scenarios, refining commands based on execution errors. Our contributions strive to transform 5G network creation, making it accessible, efficient, and responsive to diverse real-world scenarios. This research lays the foundation for a future where 5G networks become more widely attainable and adaptable to varying user needs.

References

[1] Alina Leidinger, Robert van Rooij, and Ekaterina Shutova, *The language of prompting: What linguistic properties make a prompt successful?*, arXiv preprint arXiv:2311.01967 (2023).

[2] B. A. y Arcas, *Do large language models understand us?*, Daedalus, vol. 151, no. 2, pp. 183–197, 2022.

[3] A. Chernyavskiy, D. Ilvovsky, and P. Nakov, *Transformers: "the end of history" for natural language processing?*, in Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part III 21. Springer, 2021, pp. 677–693.

[4] Bonan Min, et al., *Recent advances in natural language processing via large pre-trained language models: A survey*, ACM Computing Surveys 56.2 (2023): 1-40.

[5] Jacob Devlin Ming-Wei Chang, and Lee Kristina Toutanova, *Bert: Pre-training of deep bidirectional transformers for language understanding*, Proceedings of naacL-HLT. Vol. 1. 2019.

[6] OpenAI, *ChatGPT*, chat.openai.com.