StratoEye: An Open Platform for Language Model Integration in Air Traffic Surveillance and Optimization

Thinh Hoang[1], Huijuan Yang[2], Laurent Lapasset[1], Daniel Delahaye[1]
[1]École Nationale de l'Aviation Civile, Toulouse, France
[2]Civil Aviation Flight University of China, Sichuan, China

October 9, 2023

1 INTRODUCTION

1.1 In a Few Words

StratoEye is analogous to the Video Assistant Referee (VAR) system used in football. It monitors voice communication between controllers and pilots, analyzes trajectory data, and issues recommendations to air traffic controllers for the assessment of potentially hazardous situations within the airspace.

1.2 Background

Large Language Models (LLMs) like GPT [1] and BART [2] have demonstrated remarkable proficiency in language comprehension. It's posited that the aviation research community could harness this capability in diverse areas, including aircraft surveillance, anomaly detection, and the generation of control datasets to enhance optimization and machine learning methods, such as reinforcement learning through imitation [3]. Even with the presence of Controller Pilot Data Link Communications (CPDLC), the principal mode of communication between Air Traffic Controllers (ATCOs) and Pilots remains speech-based. Thus, the aptitude to understand and analyze language data is crucial for bolstering automation–a central objective of the SESAR Project.

The efficacy of Generative AI applications, evidenced by tangible productivity gains in areas like software engineering [4], suggests that incorporating LLMs into ATM research could emerge as a pivotal trend in the foreseeable future. To

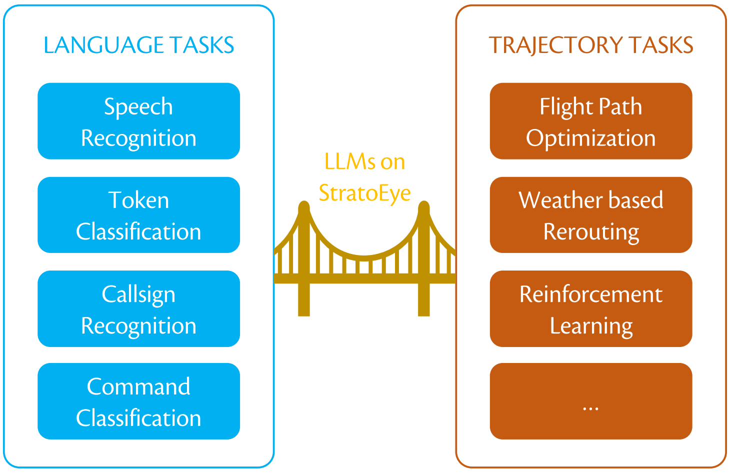

Figure 1: StratoEye aims to create a platform that allows LLMs to bridge the gap between language tasks and trajectory tasks.

our understanding, earlier studies in this vein primarily focused on speech recognition, transcription, or named-entity recognition (NER) [5–8]. These pursuits lie squarely within the domain of language modeling. There's a pressing need to expand this research trajectory to bridge the "language" and "trajectory" domains (Figure 1). By doing so, Artificial Intelligence could encompass the entirety of Air Traffic operations, even addressing tasks historically reserved for humans. Such integration could lead to noteworthy safety enhancements, such as improved surveillance capabilities, detection of misunderstood commands, and the formulation of datasets enabling machine learning algorithms to glean insights directly from human experiences.

1.3 How StratoEye can help

Three principal channels exist wherein StratoEye can provide solutions to the research quandary:

1. **Addressing the scarcity of open data.** Although one can deduce trajectory data from ADS-B packets and procure it from data aggregators like FlightRadar24 and OpenSky [9], ATC voice data remains relatively rare, non-public, costly to acquire, and demands extensive preprocessing for transcription. Moreover, ATC voice data is usually recorded separately from trajectory information, which makes it difficult to sync up. The pre-trained prowess of an LLM facilitates swift model adaptation using a limited set of examples (few-shots).
2. **Integration with trajectory processing counterparts.** As an example, to ac-

curately detect any misunderstanding in an ATCO's directive, anomaly detection algorithms require specific parameters derived solely from ATCO's speech. Conventionally, such interpretation is realized via token classification. Contrarily, StratoEye employs a generative methodology, enabling the LLM to invoke relevant trajectory processing modules using precise parameters.

3. **Decoding linguistic subtleties.** Pretrained LLMs possess the acumen to discern subtle linguistic elements like buffer words and correction endeavors, tasks that pose significant challenges to token classification approaches.

Given these advantages and StratoEye's open configuration, we anticipate a profound integration of AI within the air traffic operations. This not only augments research avenues in both the realms of language and trajectory modeling but also fosters enhanced automation. The ramifications of this could be a reduced carbon footprint for the industry, elevated passenger comfort, and augmented overall safety.

1.4 Key Components of StratoEye

StratoEye comprises of four key components:

1. A large language model based on Meta AI's LLaMA 2 [10] adapted to the ATC domain, and further fine-tuned on instructions specific to the API.
2. A Python-based aircraft simulation backend that takes various parameters in generating trajectories.
3. A web-based interface to visualize the radar tracks.
4. A unified real-time communication protocol that ties all of the above components together, allows different components to execute on different machines and infrastructures.

2 CONTRIBUTIONS

The integration of Large Language Models into trajectory processing could enhance Air Traffic Management in several ways:

1. **Safety improvement:** Large language models have the capability to understand and interpret the dialogues between Air Traffic Controllers and pilots. By simultaneously accessing radar information, these models offer an added layer of oversight in the airspace. Currently, we've incorporated an anomaly detection module based on [11] to alert air traffic controllers promptly about

potential deviations. In future iterations, integrating trajectory optimization and weather modules can help construct a knowledge graph. This would empower both air traffic controllers and pilots to query explanations and recommendations using natural language.

Metrics: We aim to showcase its effectiveness with an index named *average delay till first action to correct anomalies* T_{FA}. This measures the speed at which human operators can identify aircraft deviations from the given controller commands, both with and without the assistance of StratoEye.

2. **Misspoke commands:** StratoEye, by analyzing trajectory data, can swiftly identify inaccuracies or potentially perilous commands. This enables controllers to rectify the command before it's issued.

   Metrics: We've designed metrics to assess StratoEye's precision in classifying hazardous commands using our in-house dataset. Additionally, we measure the concordance of StratoEye's suggestions with professional ATCO judgments, gauged by the Acceptance Rate.

3. **Integration with CPDLC:** In the future, we envision that StratoEye can act as an intermediary for CPDLC, ensuring a seamless shift between text and voice links.

4. **Increased Productivity:** An anticipated outcome is that each Air Traffic Controller will be able to competently handle a greater number of aircraft within the airspace.

   Metrics: This is assessed by examining the workload, specifically, the frequency of off-nominal events per aircraft supervised, both with and without StratoEye's support.

REFERENCES

[1] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever *et al.*, "Improving language understanding by generative pre-training," 2018.

[2] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," *arXiv preprint arXiv:1910.13461*, 2019.

[3] H. Zhang and T. Yu, "Alphazero," *Deep Reinforcement Learning: Fundamentals, Research and Applications*, pp. 391–415, 2020.

[4] S. Noy and W. Zhang, "Experimental evidence on the productivity effects of generative artificial intelligence," *Available at SSRN 4375283*, 2023.

[5] J. Zuluaga-Gomez, K. Vesely, I. Szoke, P. Motlicek, M. Kocour, M. Rigault, K. Choukri, A. Prasad, S. S. Sarfjoo, I. Nigmatulina *et al.*, "Atco2 corpus: A large-scale dataset for research on automatic speech recognition and natural language understanding of air traffic control communications," *arXiv preprint arXiv:2211.04054*, 2022.

[6] D. Guo, Z. Zhang, P. Fan, J. Zhang, and B. Yang, "A context-aware language model to improve the speech recognition in air traffic control," *Aerospace*, vol. 8, no. 11, p. 348, 2021.

[7] Y. Lin, "Spoken instruction understanding in air traffic control: Challenge, technique, and application," *Aerospace*, vol. 8, no. 3, p. 65, 2021.

[8] S. Badrinath and H. Balakrishnan, "Automatic speech recognition for air traffic control communications," *Transportation research record*, vol. 2676, no. 1, pp. 798–810, 2022.

[9] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, and M. Wilhelm, "Bringing up opensky: A large-scale ads-b sensor network for research," in *IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*. IEEE, 2014, pp. 83–94.

[10] H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale *et al.*, "Llama 2: Open foundation and fine-tuned chat models," *arXiv preprint arXiv:2307.09288*, 2023.

[11] T. H. Dinh, V. Martinez, and D. Delahaye, "Recognition of outlying driving behaviors: A data-driven perspective with applications to v2x collective perception," in *2021 IEEE Vehicular Networking Conference (VNC)*. IEEE, 2021, pp. 52–59.