From Voice to Safety: Language AI Powered Pilot-ATC Communication Understanding for Airport Surface Movement Collision Risk Assessment

Yutian Pang[1] Andrew Kendall[1] Alex Porcayo[1] Mariah Barsotti[1] John-Paul Clarke[1]

[1]Department of Aerospace Engineering and Engineering Mechanics

Overview

This work introduces an integrated framework that leverages language AI to understand pilot-ATC communications, enhancing collision risk assessment during airport surface movements. Our approach combines a novel ATC rule-enhanced Named Entity Recognition (NER) model with a probabilistic risk model based on node-link airport graphs and log-normal taxi speed distributions. The methodology is validated via two case studies happened in 2024: a runway incursion at Haneda and a taxiway collision at KATL.

Motivation

Recent aviation incidents (including four major U.S. crashes in first six weeks of 2025) highlight the critical need to improve surface safety and augment existing systems. Miscommunications between pilots and controllers have contributed to runway incursions and taxiway collisions. Our framework addresses this issue by fusing advanced NLP with probabilistic risk modeling, with attention to domain-specific knowledge.

Highlights

- Novel framework integrating advanced NLP-based communication transcript analysis with surface collision risk assessment.
- Hybrid rule-enhanced NER model to extract callsigns, aircraft states, and destination intents.
- Spatiotemporal risk formulation using node-link airport layouts and log-normal speed distributions.
- Validation via two case studies (Haneda and KATL) replicating real-world airport ground movement collision scenarios.

Methodology Overview

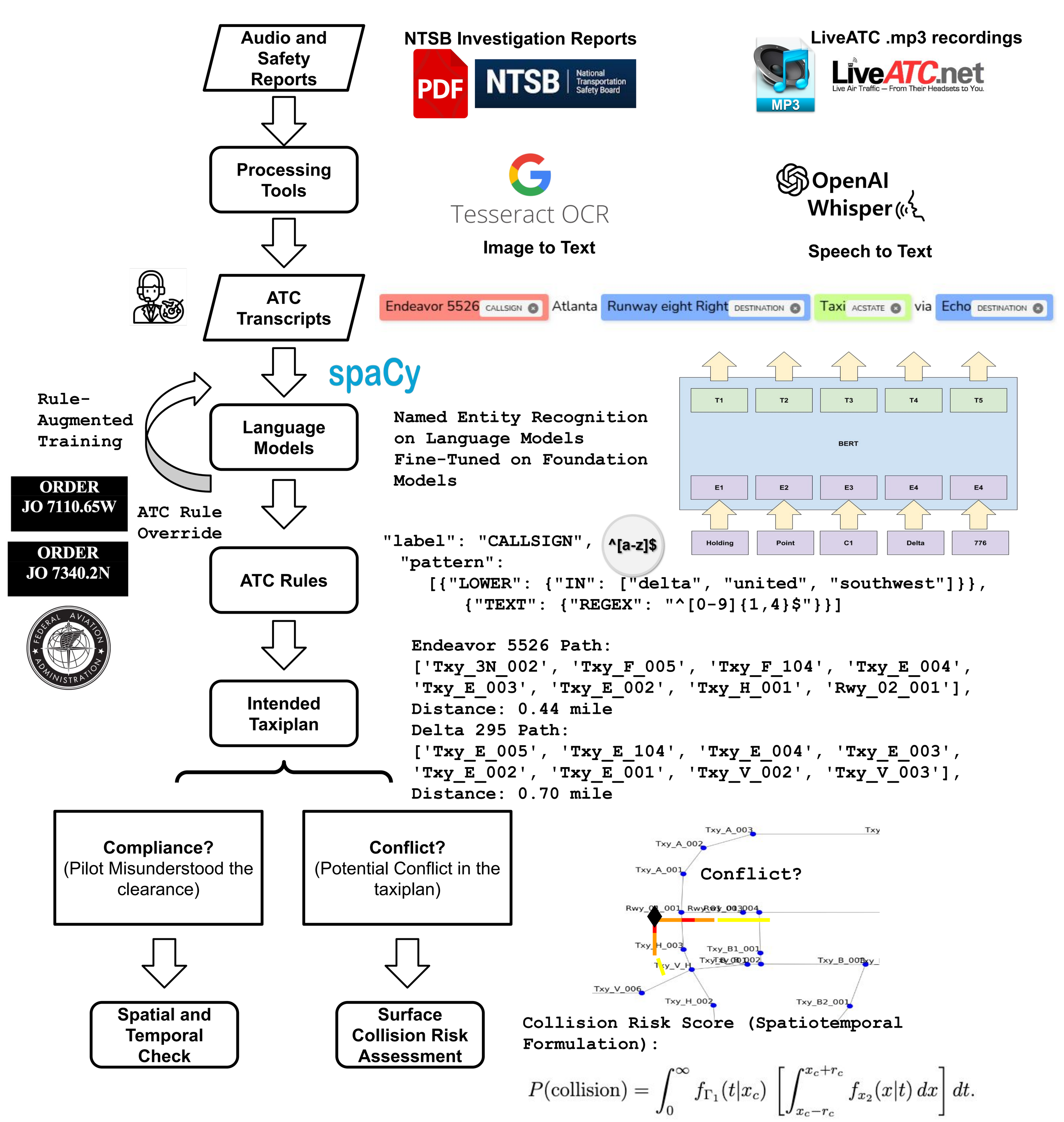


Figure 1. The proposed workflow of ATC communication transcript understanding and surface movement risk assessment.

Rule-Enhanced ATC Communication Understanding

- **Data & Annotation:** Creation of an ATC communication transcript dataset (using open-source audio, OCR of reports, etc.).
- **Rule Integration:** Incorporation of FAA-regulated phraseologies (from FAA Orders JO 7110.65W and JO 7340.2N) via SpaCy's `EntityRuler`.
- **Embedding Models:** Evaluation of seven token-level embedding models (e.g., BERT, RoBERTa, mBERT, DistilBERT, BART) shows that hybrid rule integration improves recall and overall F1 scores.

Key Findings:

1. **NER Performance:** Integrating ATC rules improves recall and F1 scores across all models.
2. **Complexity Trade-offs:** Transformer-based models yield superior accuracy but require higher computational resources.
3. **Risk Simulation:** Case study simulations (Haneda runway incursion and KATL taxiway collision) reproduce the collision events and highlight high-risk nodes.

Surface Movement Collision Risk Modeling

The collision risk modeling is based on the joint probability that two aircraft simultaneously occupy the same node in a node-link airport layout graph. The risk is computed by considering the overlap in arrival times at these nodes. The total travel time Γ_k for the k-th aircraft traveling along n taxiway links is expressed as the sum of individual link travel times:

$$\Gamma_k = \sum_{i=1}^{n} \tau_{k,i}, \quad \text{where} \quad \tau_{k,i} = \frac{d_{k,i}}{v_{k,i}}$$

Each taxi speed $v_{k,i}$ on link i is modeled as log-normal:

$$v_{k,i} \sim \text{Lognormal}(\mu_{k,i}, \sigma_{k,i}^2)$$

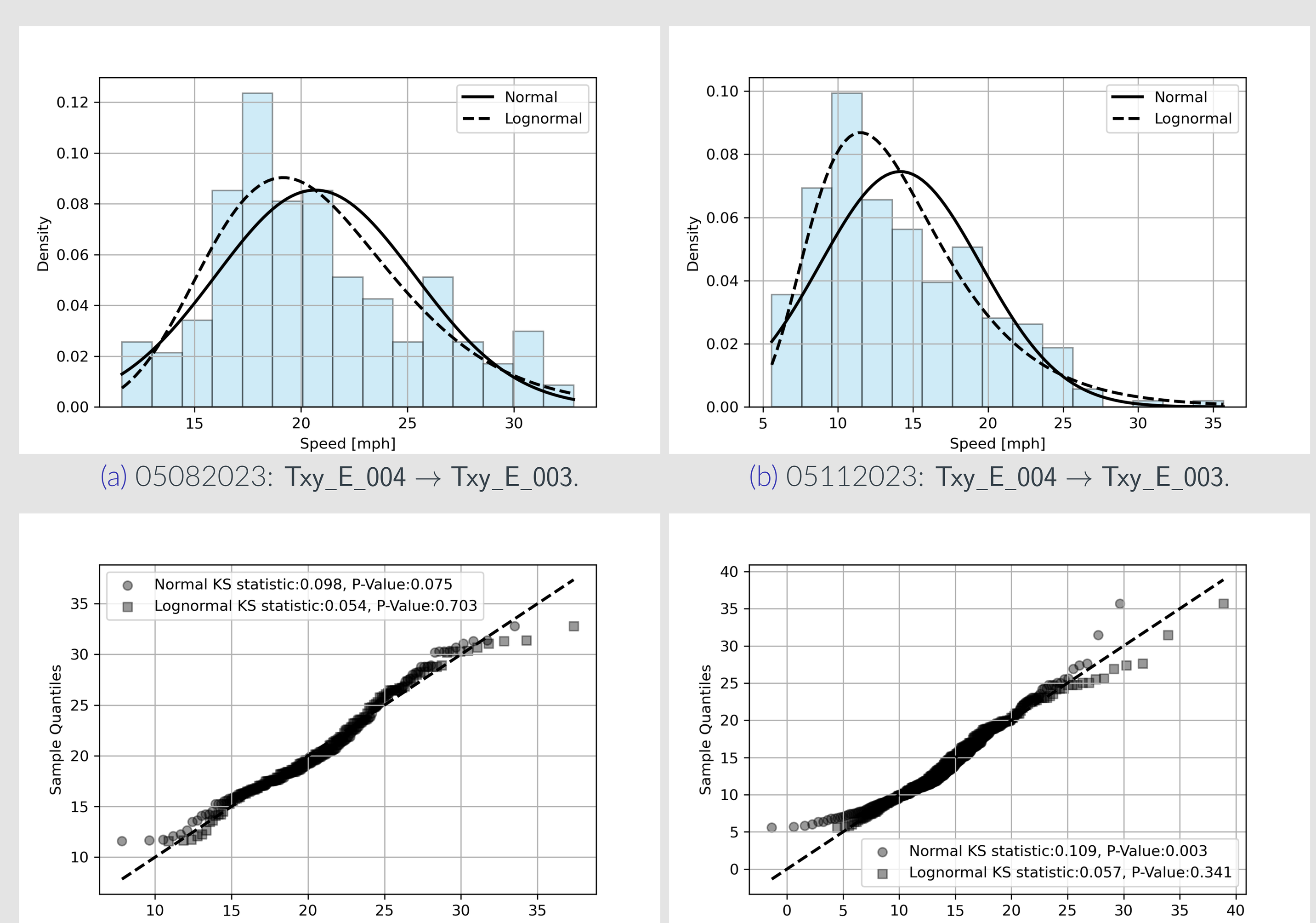
Thus, each individual link travel time $\tau_{k,i}$ is also log-normal with parameters:

$$\tau_{k,i} \sim \text{Lognormal}(\ln d_{k,i} - \mu_{k,i}, \sigma_{k,i}^2)$$

The total travel time distribution $f_{\Gamma_k}(t_k)$ is further computed as the convolution of individual link travel time distributions. A collision is defined as two aircraft occupying the same node location x_c at the same time t, within a small spatial radius r_c.

$$P(\text{collision}) = \int_0^\infty f_{\Gamma_1}(t|x_c) \left[\int_{x_c-r_c}^{x_c+r_c} f_{X_2}(x|t)\, dx\right] dt.$$

Taxi Speed Parametric Studies



(a) 05082023: Txy_E_004 → Txy_E_003. (b) 05112023: Txy_E_004 → Txy_E_003.

(c) 05082023: Txy_E_004 → Txy_E_003. (d) 05112023: Txy_E_004 → Txy_E_003.

Figure 2. **Log-normal** distributions are better fits in these cases.

Case Studies

Table 1. Key ATC communication transcript extracted with the knowledge-enhanced hybrid learning model for the 2024 KATL taxiway collision case study.

TIME	CALLSIGN	AC_STATE	DEST_RUNWAY	DESTINATION
0:08	Delta 295	taxi	08R	Romeo
0:14	Delta 295	taxi	08R	Rwy_02_001
0:20	Delta 295	Taxi	08R	foxtrot
0:33	Delta 295	continue,hold	08R	ramp 5
0:44	Delta 295	give way,inbound,join	08R	Echo(Txy_E_002)
0:50	Delta 295	give way	08R	
0:57	Endeavor 5526	taxi	08R	Rwy_02_001
1:27	Delta 295	go	08R	
1:35	Delta 295	continue,hold	08R	
1:45	Delta 295	holding	08R	Victor(Txy_V_003)
1:54	Endeavor 5526	line up,wait	08R	
2:10	Endeavor 5526	collision		
2:10	Delta 295	collision		

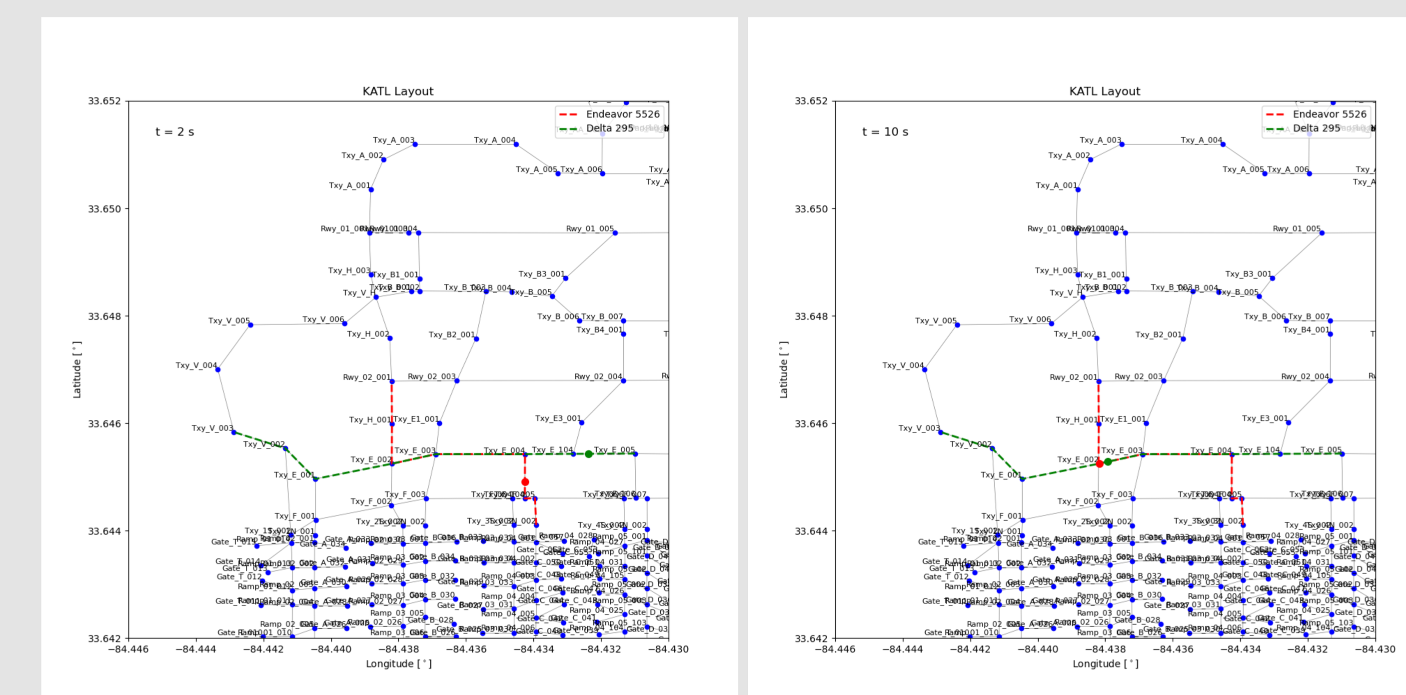
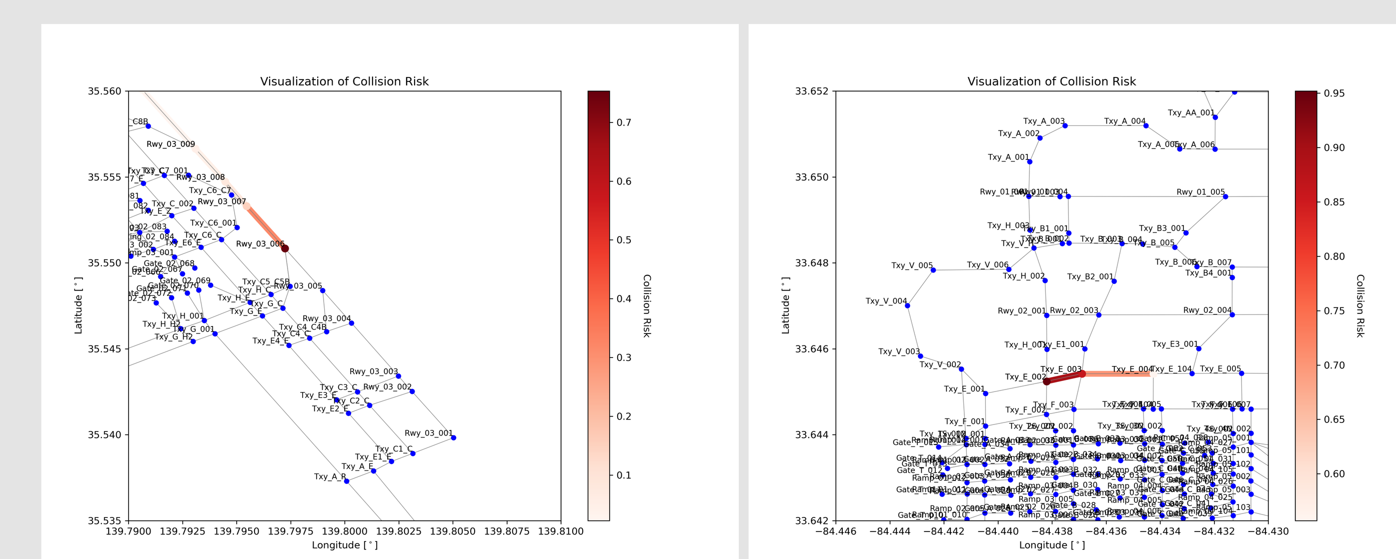


Figure 3. Node-link simulation of the accident happened at the Henada Airport in January 2024.



(a) Risk score for the 2024 Henada runway collision scenario. (b) Risk score for the 2024 KATL taxiway collision scenario.

Figure 4. Collision risk score visualized by color on the node-link graph. The scores are evaluated at the overlapping node links from the generated path, which are generated based on the language model learned destination intents.

Conclusions

- The integrated framework effectively fuses language AI and risk modeling to enhance airport surface safety.
- Incorporating domain-specific ATC rules into NER significantly boosts extraction accuracy.
- Simulations validate the approach by accurately predicting high-risk collision nodes.

Future Work

- Extend the model to multi-aircraft collision scenarios.
- Fuse computer vision with NLP for real-time safety monitoring.
- Enhance model security using differential privacy for embeddings.