

Terminal Area Traffic Situation Prediction using a Transformer-based Model with Multi-Head Self-Attention

Aniket Pal¹ and Nisha Wandile²

¹⁻²Department of Computer Engineering, JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, 411033, India
Email: aniketpal.edu@gmail.com, nishakimmatkar@gmail.com

Abstract— Terminal areas such as airports, ports, and logistics centers often face complex traffic conditions due to continuous vehicle flow and unpredictable patterns. Addressing these issues requires accurate and timely forecasting to reduce congestion, enhance safety, and ensure efficient resource utilization. This study introduces a forecasting model based on the Multi-Head Attention Transformer architecture, which leverages self-attention mechanisms to identify complex temporal relationships in traffic data. By using multiple attention heads, the model captures patterns across different time scales, improving the prediction of both short-term variations and long-term trends. When evaluated against traditional models like ARIMA, LSTM, and GRU, the proposed approach demonstrated notably higher accuracy. While the model shows strong potential for real-time traffic management in various terminal settings, it also demands extensive data and computational power to function optimally.

Index Terms— Traffic situation prediction, transformer, temporal convolutional network, Intelligent transportation system, air transportation.

I. INTRODUCTION

Traditional air traffic management (ATM) systems face challenges due to the increasing complexity of airport operations. To address this, Intelligent Transportation Systems (ITS) have emerged, leveraging AI to improve efficiency and safety. Smart ATM relies on situational awareness—understanding current conditions and predicting future trends. While prior research using network theory and deep learning has explored broader airspace complexity, terminal area prediction remains underexplored. To fill this gap, this study introduces the ConvTrans-TCN model, which processes time-series traffic data through encoding, synthesis, and prediction stages. Its accurate forecasts can support improved decision-making in terminal traffic management.

II. RELATED WORK

The increasing complexity of air traffic has outpaced traditional Air Traffic Management (ATM) systems, prompting researchers to adopt Intelligent Transportation Systems (ITS) for improved situational awareness and traffic optimization [1], [2]. Early studies used complex network theory to model air traffic complexity and congestion [3]–[6]. More recently, deep learning models like spatiotemporal hybrids [7] and graph convolutional networks [8] have been applied, though they often overlook terminal areas, which are the most congested.

Transformer models, particularly those using multi-head self-attention [20], have shown promise in capturing long-term dependencies in time-series data. Enhancements like GMAN [18], NAST [19], and Bidirectional Transformers [17] improve spatial-temporal modeling but still struggle with localized feature extraction in dense environments. To address these limitations, temporal convolutional networks (TCNs) and causal convolutions have been integrated with attention mechanisms to improve local and long-range pattern recognition [22], [24]. However, few models specifically focus on terminal area traffic prediction. This study proposes a ConvTrans-TCN model, which combines multi-head attention, causal convolution, and TCNs to effectively model both short- and long-term dependencies for real-time ATM support.

III. PROPOSED METHODOLOGY

A. Transformer

Introduced by Google in 2017, the Transformer model addressed RNN limitations in capturing long-range dependencies. This study uses only the Encoder, leveraging self-attention and feedforward layers to process sequences without accessing future data.

Multi head and self attention mechanism

Self-attention captures sequence relationships using query, key, and value vectors. Multi-head attention improves this by combining several attention heads to extract richer context, followed by a linear transformation. At the outset, the query grid $Q \in \mathbb{R}^T \times D_q$, the key of the matrix $K \in \mathbb{R}^T \times D_k$, and the resulting matrix $V \in \mathbb{R}^T \times D_v$ are all created by linearly transforming the input matrix $X \in \mathbb{R}^T \times D_x$. Next, an attention-scoring function is used to calculate the attention matrix $MA \in \mathbb{R}^T \times T$ [26]. The attention balance matrix $WA \in \mathbb{R}^T \times T$ is generated by normalizing the result with the softmax function. Finally, the value matrix and the weight for attention matrix are multiplied by each other to yield the output $H \in \mathbb{R}^T \times D_v$.

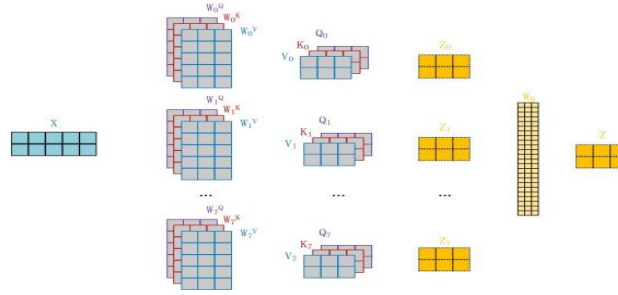


Figure 1: System Architecture

Position by position and feed forward networks

The feedforward network includes two dense layers separated by a ReLU activation. It processes the encoded input using weight matrices W_1 , W_2 and biases b_1 , b_2 to refine and integrate information.

$$F_{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (8)$$

Enhanced transformer model

Terminal traffic data, as a time-series, requires accurate modeling of both short- and long-term trends while managing anomalies. Zhou et al. [22] proposed a causal convolutional self-attention module that improves local pattern detection. Unlike standard attention mechanisms, it uses a Conv1D layer with causal padding to maintain sequence order and prevent information leakage. A triangular mask is also applied to the attention matrix to enhance prediction accuracy.

$$Q = \text{Conv1D}(X) \quad (9)$$

$$K = \text{Conv1D}(X) \quad (10)$$

$$H = \text{softmax}(QK^T / \sqrt{D_K} + \text{Mask})V \quad (11)$$

B. Information fusion based model on tcn

To predict terminal-area traffic, a causal convolutional self-attention module encodes the input sequence into feature representations. This is followed by a Temporal Convolutional Network (TCN) for data fusion and final

prediction [23]. The TCN maintains sequence order using stacked 1D causal convolutions and integrates Residual Connections with Dilated Convolution to capture both short- and long-range dependencies efficiently [24], [26].

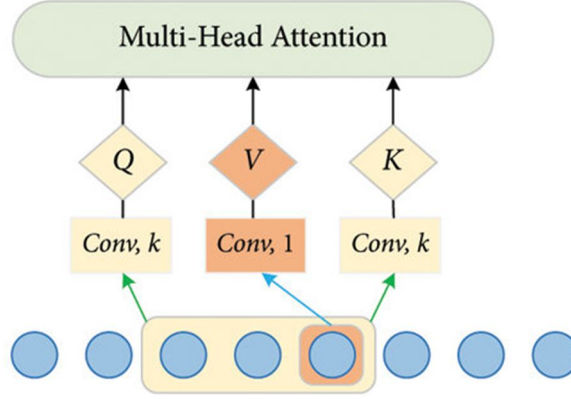


Figure 2. Multi head causal convolutional self attention

Dilated causal convolution

Dilated causal convolution expands the model's receptive field exponentially by increasing the dilation rate with network depth [26]. This allows the network to capture long-range dependencies more efficiently than simply increasing kernel size, enabling better temporal feature fusion and improved sequence modeling (see Fig. 4)

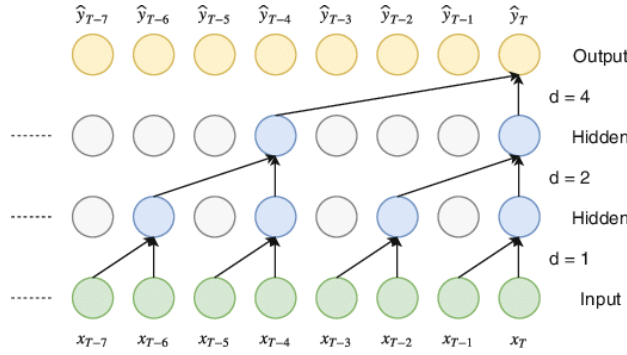


Figure 3. Dilated causal convolution

Residual connection

Residual connections add the original input directly to the output of a layer's transformation function, helping deeper networks learn effectively and avoid vanishing gradients. To align input and output dimensions, a 1D convolution with kernel size 1 is applied [26].

C. ConvTrans-TCN-based situation prediction model

The ConvTrans-TCN model improves terminal traffic prediction through feature extraction, temporal fusion, and scenario estimation. It uses a causal convolutional self-attention layer, followed by fully connected layers with ReLU, residuals, and normalization. A TCN with dilated convolutions captures temporal patterns, and the final output is derived from the sequence's last step. MSE is used as the loss function, ensuring strong generalization through regularization.

IV. RESULTS AND DISCUSSION

A. Preparation transformer

The study uses 10-minute interval traffic data from Tianjin Binhai Airport (June 3–16, 2019), including 12 features and one labeled status. Data was normalized, split (70% training, 10% validation, 20% testing), and expanded using interpolation with a 10-step sliding window. The model combines a causal convolutional encoder, 4-head attention, and a 6-layer dilated TCN. It was trained using the Adam optimizer and evaluated with RMSE and MAE metrics.

$$\text{RMSE}(X, f) = \sqrt{\frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2} \quad (22)$$

$$\text{MAE}(X, f) = \frac{1}{N} \sum_{i=1}^N |f(x_i) - y_i| \quad (23)$$

B. Predictability Analysis of Terminal Area Traffic

Short input sequences often lack sufficient data, causing prediction uncertainty, while longer sequences raise computational costs. Therefore, finding an optimal sequence length is essential. Using Liu et al.'s entropy estimation method [26], the study evaluated the bounds of traffic predictability. Results show that beyond a length of 2000, entropy and predictability stabilize—indicating that the dataset has consistent patterns and is suitable for reliable forecasting

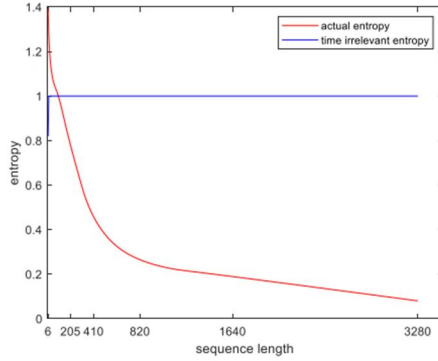


Figure 4: Relationship between Sequence Length and Entropy

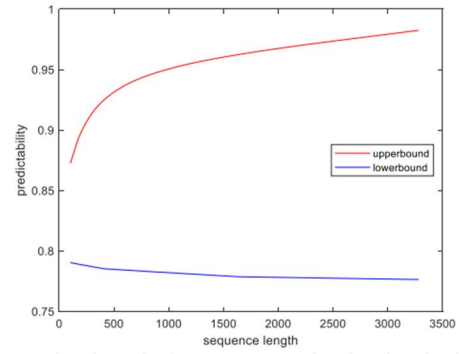


Figure 5: Relationship between sequence length and traffic situation predictability

C. Effects of different parameters to the proposed model

The study evaluates TCN, Trans-TCN, and ConvTrans-TCN models for terminal traffic prediction. ConvTrans-TCN performs best, combining causal convolution and multi-head attention to capture both short- and long-term dependencies. Causal padding further improves accuracy. Optimal results were achieved with 4 attention heads and encoder-decoder layers. Compared to LSTM, GA-GMNN, and BP, ConvTrans-TCN shows superior predictive accuracy and generalization. Visualization techniques (PCA, t-SNE) and residual analysis validate the model's robustness for proactive air traffic management.

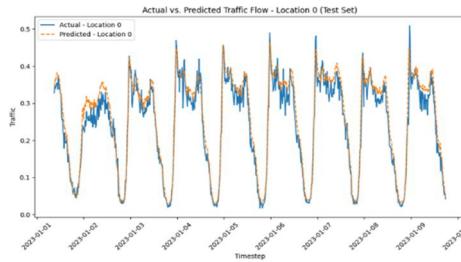


Figure 6: PCA Scatter Plot of First Two Components

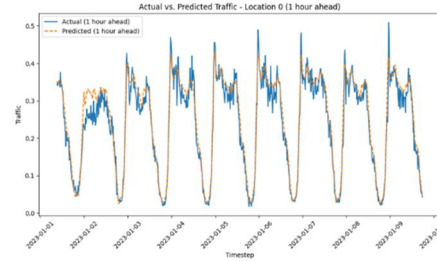


Figure 7: t-SNE Scatter Plot

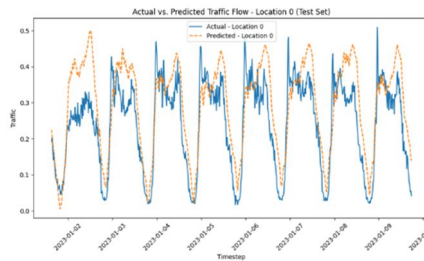


Figure 8: Predicted vs Actual Traffic Flow

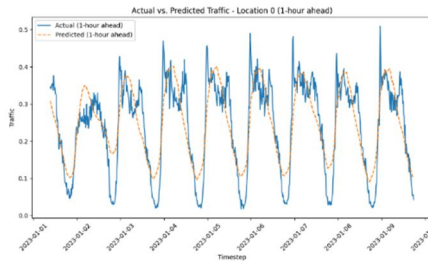


Figure 9: Residuals Plot

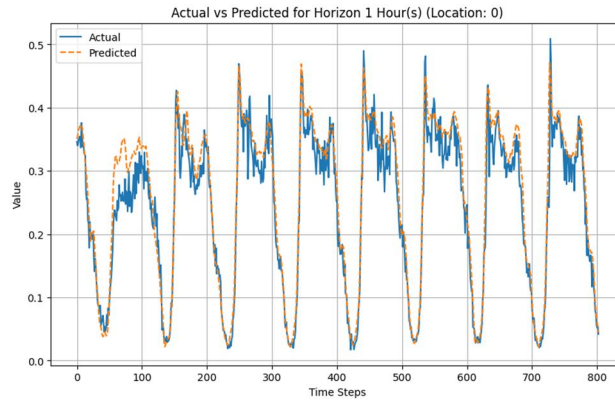


Figure 10: Feature Importance

V. CONCLUSIONS

Accurate terminal traffic forecasting is key to efficient air traffic management. This study presents a transformer-based model with multi-head attention and temporal layers. Compared to LSTM, BP, and GA-GMNN, it achieves better accuracy and lower MAE.

Its parallel processing supports real-time use. Future improvements may include external factors like weather and application across other terminals.

REFERENCES

- [1] K.Golestan,R.Soua,F.Karray,andM.S.Kamel,“Situationawareness within the context of connected cars: A comprehensive review and recent trends,” *Inf. Fusion*, vol. 29, pp. 68–83, May 2016.
- [2] M. R. Endsley, “Design and evaluation for situation awareness enhance- ment,” in *Proc. Hum. Factors Soc. Annu. Meeting*, vol. 32. Los Angeles, CA, USA: Sage Publications, 1988, p. 2.
- [3] M.Prandini,L.Piroddi,S.Puechmorel,andS.L.Brazdilova,“Towardair traffic complexity assessment in new generation air traffic management systems,” *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 809–818, Sep. 2011.
- [4] W.Hongyong,S.Ziqi,andW.Ruiying,“Studyonevolutioncharacteristics of air traffic situation complexity based on complex network theory,” *Aerosp. Sci. Technol.*, vol. 58, pp. 518–528, Nov. 2016.
- [5] H.Chen,“Airtrafficcomplexityassessmentbasedonordereddeepmet- ric,” *Aerosp. Sci. Technol.*, vol. 9, no. 12, pp. 518–528, Nov. 2022.
- [6] Y. Kawagoe, R. Chino, S. Tsuzuki, E. Itoh, and T. Okabe, “Analyzing stochastic features in airport surface traffic flow using cellular automaton: Tokyo international airport,” *IEEE Access*, vol. 10, pp. 95344–95355, 2022.
- [7] W. Du, B. Li, J. Chen, Y. Lv, and Y. Li, “A spatiotemporal hybrid model for airspace complexity prediction,” *IEEE Intell. Transp. Syst. Mag.*, early access, Sep. 28, 2022, doi: 10.1109/MITS.2022.3204099.
- [8] D. Sui, K. Liu, and Q. Li, “Dynamic prediction of air traffic situation in large-scale airspace,” *Aerospace*, vol. 9, no. 10, p. 568, Sep. 2022.
- [9] H. J. Park, T. Kim, Y. S. Kim, J. Min, K. W. Sung, and S. W. Han, “CRFormer: Complementary reliability perspective transformer for auto- motive components reliability prediction based on claim data,” *IEEE Access*, vol. 10, pp. 88457–88468, 2022.
- [10] C.Wang,Y.Chen,S.Zhang,andQ.Zhang,“Stockmarketindexprediction using deep transformer model,” *Expert Syst. Appl.*, vol. 208, Dec. 2022, Art. no. 118128.
- [11] L.Wu,Y.Wang,X.Li,andJ.Gao,“Deepattention-basedspatiallyrecur- sive networks for fine-grained visual recognition,” *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1791–1802, May 2019.
- [12] L. Wu, Y. Wang, J. Gao, M. Wang, Z.-J. Zha, and D. Tao, “Deep coattention-based comparator for relative representation learning in person re-identification,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 722–735, Feb. 2021.
- [13] S.-J. Bu and S.-B. Cho, “Time series forecasting with multi-headed attention-based deep learning for residential energy consumption,” *Ener- gies*, vol. 13, no. 18, p. 4722, Sep. 2020.
- [14] W.Fang,W.Zhuo,J.Yan,Y.Song,D.Jiang,andT.Zhou,“Attentionmeets long short-term memory: A deep learning network for traffic flow forecast- ing,” *Phys. A, Stat. Mech. Appl.*, vol. 587, Feb. 2022, Art. no. 126485.
- [15] R.Huang,C.Huang,Y.Liu,G.Dai,andW.Kong,“LSGCN:Longshort- term traffic prediction with graph convolutional networks,” in *Proc. 29th Int. Joint Conf. Artif. Intell.*, vol. 7, Jul. 2020, pp. 2355–2361.
- [16] X. Kong, J. Zhang, X. Wei, W. Xing, and W. Lu, “Adaptive spatial–temporal graph attention networks for traffic flow forecasting,” *Appl. Intell.*, vol. 52, no. 4, pp. 4300–4316, Mar. 2022.

- [17] C. Chen, Y. Liu, L. Chen, and C. Zhang, “Bidirectional spatial–temporal adaptive transformer for urban traffic flow forecasting,” *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 30, 2022, doi: 10.1109/TNNLS.2022.3183903.
- [18] C. Zheng, X. Fan, C. Wang, and J. Qi, “GMAN: A graph multi-attention network for traffic prediction,” in *Proc. AAAI Conf. Artif. Intell.*, Palo Alto, CA, USA, 2020, vol. 34, no. 1, pp. 1234–1241.
- [19] K.Chen,G.Chen,D.Xu,L.Zhang,Y.Huang,andA.Knoll,“NAST:Non- autoregressive spatial–temporal transformer for time series forecasting,” 2021, arXiv:2102.05624.
- [20] T. Qi, G. Li, L. Chen, and Y. Xue, “ADGCN: An asynchronous dilation graph convolutional network for traffic flow prediction,” *IEEE Internet Things J.*, vol. 9, no. 5, pp. 4001–4014, Mar. 2022.
- [21] Wandile, N., Bhonde, A., Garje, A., Chavan, P., & Bhangale, P. Blockchain-based Integration of Social Media and Music for Enhanced Privacy and Trust in a Decentralized Ecosystem. *Grenze International Journal of Engineering & Technology (GIJET)*, 10, 2024.