<u>ACK 2024 학술발표대회 논문집 (31권 2호)</u>
TSANTP: 공간 코딩 주의 메커니즘을 통합한 새로운
- 네트워크 트래픽 예측 모델 <u>ACK 2024학술발표대회 논문집 (31권 2호)</u>
간 코딩 주의 메커니즘을 통합한 새로운
네트워크 트래픽 예측 모델
이용비 !, 김경백 ²
'전남대학교 인공지능융합학과 석사과정 ACK 2024 학술발표대회 논문집 (31권 2호)
1**년 국일 대거니즘을 통합한 새로운**
제트워크 트래픽 예측 모델
이용비트 지경백 ²
이용비트 인공지능용합학과 석사과정
^{2전남대학교} 인공지능용합학과 교수 <u>ACK 2024 학술발표대회 논문집 (31권 2호)</u>
코딩 주의 메커니즘을 통합한 새로운
트워크 트래픽 예측 모델
이용비 ! 김경백 2
일남대학교 인공지능융합학과 석사과정
²전남대학교 인공지능융합학과 교수
₂(gfei0712@gmail.com, kyungbaekkim@jnu.ac.kr 간 코딩 주의 메커니즘을 통합한 시
네트워크 트래픽 예측 모델
- 이용비트 김경백?
- ^{이용비트 김경백?
- ^{이용비트 김경박 석사과정
- ²전남대학교 인공지능융합학과 교수
lilongfei0712@gmail.com, kyungbaekkim@jnu.ac.kr
DVel Network Traffic Prediction Mo}} **TSANTP: 공간 코딩 주의 메커니즘을 통합한 새로운**
비트워크 트래픽 예측 모델
^{이용비 , 김경맥?
^{이용비 ,} 김경맥?
^{이용비 ,} 김경맥?
^{이용비 ,} 김정맥?
^{이용비 ,} 김정맥?
^{이용비 ,} 김정맥?
^{이용비 ,} 김정맥?
^{이용비 서}작적
¹ 김남대학교 인종지능용합학과 교수
I**TSANTP: A Novel Network Traffic Prediction Model} eCK 2024 학술발표대회 논문집 ©1권 2호)**
SANTP: 공간 코딩 주의 메커니즘을 통합한 새로운
- 네트워크 트래픽 예측 모델
- ^{이용비} 김경_백을
- ^{이용비} 김경백^을
- 결담대학교 인공지능용합학과 교수
- ¹전담대학교 인공지능용합학과 교수
- Illongtion Legrating Space Coding and Attention Mechanisms
- LongFei

이용비 ¹ , 김경백 ²

이용비 ', 김경백 ²
'대학교 인공지능융합학과 석사과정
'남대학교 인공지능융합학과 교수
i0712@gmail.com, kyungbaekkim@jnu.ac.kr
**Network Traffic Prediction
Coding and Attention Mec**
LongFei Li ¹, Kyungbaek Kim ²
telligence Convergence, Chonnam National 이용비 !, 김경백 2
'전남대학교 인공지능용합학과 석사과정
²전남대학교 인공지능용합학과 교수
lilongfei0712@gmail.com, kyungbaekkim@jnu.ac.kr
FP: A Novel Network Traffic Prediction Model Int
Artificial Intelligence Convergence, Chonnam National University il.com, kyungbaekkim@jnu.ac.kr
 Ork Traffic Prediction Model Int
 g and Attention Mechanisms

 $\frac{1}{2}$, Kyungbaek Kim²

Convergence, Chonnam National University

 Q e^t

ethnology, network traffic has increa

MIP: A Novel Network Traffic Prediction Model Int

egrating Space Coding and Attention Mechanisms

LongFei Li¹, Kyungbaek Kim²

Dept. of Artificial Intelligence Convergence, Chonnam National University
 $\frac{Q}{2}$ et
 Example 18 Space Coding and Attention Mechanisms

LongFei Li⁻¹, Kyungbaek Kim⁻²

Dept. of Artificial Intelligence Convergence, Chonnam National University
 Ω

With the widespread application of 5G technology, net **Example 18 Space Coding and Attention Mechanisms**

LongFei Li¹, Kyungback Kim²

Dept. of Artificial Intelligence Convergence, Chonnam National University

With the widespread application of 5G technology, network tra **Example 19 and Example Country and Example Convolutional networks (GCN)**
 Example 19 and 19 a Example 1. I., Syungback Kim ²

Dept. of Artificial Intelligence Convergence, Chonnam National University

With the widespread application of 5G technology, network traffic has increased unprecedentedly, which

the ha LongFei Li¹, Kyungbaek Kim²

Dept. of Artificial Intelligence Convergence, Chonnam National University
 Ω

With the widespread application of 5G technology, network traffic has increased unprecedentedly, which has Dept. of Artificial Intelligence Convergence, Chonnam National University

With the widespread application of 5G technology, network traffic has increased unprecedentedly, which has a significant impact on network traffic With the widespread applice

ch has a significant impact on 1

rely on time series analysis of

ic. Graph convolutional network

models, spatiotemporal features

s paper introduces a new netwo

which aims to more effective Ω \cong
With the widespread application of 5G technology, network traffic has increase
ch has a significant impact on network traffic management. Traditional network tra
rely on time series analysis of seasonal parters 22 V

22 V

24 V

24 V

24 C

24 C

24 C

24 C

24 C

24 C

26 C

36 C

37 C

37 C

37 C

39 C
 With the widespread application of 5G technology, network traffic has increased unp

ch has a significant impact on network traffic management. Traditional network traffic p

rely on time series analysis of seasonal patter exhibiting a significant impact on network traffic management. Traditional network traffic
rely on time series analysis of seasonal patterns, ignoring the inherent spatial correlation
i.e. Graph convolutional networks (GCN **24**

network traffic has increased unprecedentedly, whi

nent. Traditional network traffic prediction methods

rig the inherent spatial correlation of network traff

correlations. By combining GCN with time series

necous

rely on time series analysis of seasonal patterns, ignoring the inherent spatial correlation

i. Graph convolutional network (GCN) learn spatial correlations. By combining GCN

models, spaticulary factures can be captured ic. Graph convolutional networks (GCN) learn spatial correlations. By combining GCN
models, spatiotemporal features can be captured simultaneously, thereby improving predic
s paper introduces a new network traffic predicti models, spatotemporal reatures can be captured simultaneously, thereby improving pre

s paper introduces a new network traffic prediction model TSA-NTP based on the a

which aims to more effectively capture spatiotemporal s paper into the set and we here the prediction indeed 15A-1817 based of which aims to more effectively capture spatiotemporal features in complex net

1. **Introduction** methods are increasin

with the rapid increase in th **1. Introduction**
 1. Introduction

With the rapid increase in the number of mobile devices

and connections globally, particularly with the widespread

greater efficacy in extra

adoption of 5G technology, global networ **1. Introduction**

with the rapid increase in the number of mobile devices

and connections globally, particularly with the widespread

greater efficacy in extracting to

and connections globally, particularly with the wid **Example 1. Introduction**
 Analysis of seasonal patterns and connections globally, particularly with the widespread growth as Recurrent
 Analysis of sections and connections will constitute 15% of
 Analysis of series With the rapid increase in the number of mobile devices

and connections globally, particularly with the widespread

areas are streament increases. [4]For

exhibiting unprecedented growth. Recent data indicates that

by 20 and connections globally, particularly with the widespread
adoption of 5G technology, global network traffic is flow sequences. [4]For
exhibiting unprecedented growth. Recent data indicates that
by 2023, 5G devices and con adoption of 5G technology, global network traffic is now sequences. [4] ror in exhibiting unprecedented growth. Recent data indicates that Networks can automatical
by 2023, 5G devices and connections will constitute 15% of ibiting unprecedented growth. Recent data indicates that

2023, 5G devices and connections will constitute 15% of

10% and inspire these advancement

behal mobile devices and connections, an increase from the

primarily em by 2023, 5G devices and connections will constitute 15% of

global mobile devices and connections, an increase from the

previously projected 10%. This trend has resulted in a sharp

previously emplatize the ter-

rise in global mobile devices and connections, an increase from the
previously projected 10%. This trend has resulted in a sharp
reviously emphasize the
rise in traffic between network devices, significantly welvook the spatial co previously projected 10%. This trend has resulted in a sharp

rinnarity emphasize the t

rise in traffic between network devices, significantly overlook the spatial correlat

complicating network traffic management. Tradit

rise in traffic between network devices, significantly overlook the spatial correlations
complicating network traffic management. Traditional Metwork traffic is exchanged
network traffic prediction methods predominantly re complicating network traffic management. Traditional Network traffic is exchanged betwork traffic prediction methods predominantly rely on the antiverses network links. Due the nonlinearity and dynamic variability and extr network traffic prediction methods predominantly rely on the

analysis of seasonal patterns and time-series characteristics.

[1]However, the nonlinearity and dynamic variability behavior between these li

exhibited by rea analysis of seasonal patterns and time-series characteristics.

[1] However, the nonlinearity and dynamic variability behavior between these links. For

exhibited by real-world network traffic present substantial congeste [1] However, the nonlinearity and dynamic variability behavior between these lines, the exhibited by real-world network traffic present substantial congested links are more l
challenges to achieving high-precision predict

94

network traffic has increased unprecedentedly, whi

nent. Traditional network traffic prediction methods

correlations. By combining GCN with time series

neously, thereby improving prediction accuracy. Thi

odel TSA 2⁴
network traffic has increased unprecedentedly, whi
ment. Traditional network traffic prediction methods
ing the inherent spatial correlation of network traff
correlations. By combining GCN with time series
neously, th network traffic has increased unprecedentedly, whi
ment. Traditional network traffic prediction methods
ring the inherent spatial correlation of network traff
correlations. By combining GCN with time series
necously, there ment. Traditional network traffic prediction methods
ing the inherent spatial correlation of network traff
correlations. By combining GCN with time series
neously, thereby improving prediction accuracy. Thi
del TSA-NTP bas ing the inherent spatial correlation of network traff
correlations. By combining GCN with time series
neously, thereby improving prediction accuracy. Thi
del TSA-NTP based on the attention mechanism,
features in complex ne relations. By combining GCN with time series
susly, thereby improving prediction accuracy. Thi
1 TSA-NTP based on the attention mechanism,
tures in complex network environments.
thods are increasingly applied in various fi neously, thereby improving prediction accuracy. Thi
del TSA-NTP based on the attention mechanism,
features in complex network environments.
methods are increasingly applied in various fields, methods
such as Recurrent Neur

ovel TSA-NTP based on the attention mechanism,
features in complex network environments.
methods are increasingly applied in various fields, methods
such as Recurrent Neural Networks (RNN) have shown
greater efficacy in ex reatures in complex network environments.

methods are increasingly applied in various fields, methods

such as Recurrent Neural Networks (RNN) have shown

greater efficacy in extracting temporal features from traffic

flo methods are increasingly applied in various fields, methods
such as Recurrent Neural Networks (RNN) have shown
greater efficacy in extracting temporal features from traffic
flow sequences. [4]For instance, Convolutional Ne methods are increasingly applied in various fields, methods
such as Recurrent Neural Networks (RNN) have shown
greater efficacy in extracting temporal features from traffic
flow sequences. [4]For instance, Convolutional Ne such as Recurrent Neural Networks (RNN) have shown
greater efficacy in extracting temporal features from traffic
flow sequences. [4]For instance, Convolutional Neural
Networks can automatically extract features from networ greater efficacy in extracting temporal features from traffic
flow sequences. [4]For instance, Convolutional Neural
Networks can automatically extract features from network
traffic data, resulting in improved predictive pe flow sequences. [4]For instance, Convolutional Neural
Networks can automatically extract features from network
traffic data, resulting in improved predictive performance.
Despite these advancements, existing prediction met Networks can automatically extract features from network
traffic data, resulting in improved predictive performance.
Despite these advancements, existing prediction methods
primarily emphasize the temporal dimension and of traffic data, resulting in improved predictive performance.

Despite these advancements, existing prediction methods

primarily emphasize the temporal dimension and often

overlook the spatial correlations present in netwo Despite these advancements, existing prediction methods
primarily emphasize the temporal dimension and often
overlook the spatial correlations present in network traffic.
Network traffic is exchanged between nodes at multi marily emphasize the temporal dimension and often
erlook the spatial correlations present in network traffic.
twork traffic is exchanged between nodes at multiple sites
d traverses network links. Due to the adjacency inher overlook the spatial correlations present in network traffic.
Network traffic is exchanged between nodes at multiple sites
and traverses network links. Due to the adjacency inherent in
network topology, there is significan Network traffic is exchanged between nodes at multiple sites
and traverses network links. Due to the adjacency inherent in
network topology, there is significant correlation in traffic
behavior between these links. For exa and traverses network links. Due to the adjacency inherent in
network topology, there is significant correlation in traffic
behavior between these links. For example, links adjacent to
congested links are more likely to be network topology, there is significant correlation in traffic
behavior between these links. For example, links adjacent to
congested links are more likely to be affected, leading to the
propagation of congestion. [5] As a

Figure 1

Figure 1

capturing the complex relationships between netwo

While time-series models excel at extracting tem

spatial features, combining these techniques wi

allows for the simultaneous capture of spatic

featu Figure 1: TSANTP model structure

Moreover, with the complex relationships between network links. Where W_1 is the weight mat

in the series models excel at extracting temporal and This step enhances the model's

tital **Example 1: TSANTP model structure**
 Example 2: TSANTP model structure

While time-series models excel at extracting temporal and This step enhances the model

and features, combining these techniques with GCN

and feat **Figure 1: TSANTP model structure**

capturing the complex relationships between network links. Where W_1 is the weight

While time-series models excel at extracting temporal and This step enhances the n

spatial feature **Figure 1: TSA**
capturing the complex relationships between network lin
While time-series models excel at extracting temporal a
spatial features, combining these techniques with GC
allows for the simultaneous capture of sp but the complex relationships between network links. Where W_1 is the weight
in the secondary and this step enhances then it at features, combining these techniques with GCN
temporal features.
this paper, we see the sim capturing the complex relationships between network links. where W_1 is the we While time-series models excel at extracting temporal and This step enhances the spatial features, combining these techniques with GCN tempo

While time-series models excel at extracting temporal and

spatial features, combining these techniques with GCN

enormoral features in network traffic, thus enhancing both prediction

features in network traffic, thus en spatial features, combining these techniques with GCN

allows for the simultaneous capture of spatiotemporal

allows for the simultaneous capture of spatiotemporal

features in network traffic, thus enhancing both predict construction model, and the simultaneous capture of spatiotemporal

features in network traffic, thus enhancing both prediction
 $H_{i+1} = Re LU (Conv1D)$

Moreover, with the growing success of Transformers in

the time-series dom features in network traffic, thus enhancing both prediction
accuracy and robustness.
accuracy and robustness.
Moreover, with the growing success of Transformers in
the time-series domain, they have been incorporated into accuracy and robustness.

Moreover, with the growing success of Transformers in

the time-series domain, they have been incorporated into

the one-dimensional convolution

network traffic prediction models to further impr Moreover, with the growing success of Transformers in

the time-series domain, they have been incorporated into

the one-dimensional con-

network traffic prediction models to further improve

through multiple convolution the time-series domain, they have been incorporated into

metwork traffic prediction models to further improve

predictive accuracy. [7]

In this paper, we propose a novel attention-based network

trough multiple convectio network traffic prediction models to further improve

predictive accuracy. [7]

In this paper, we propose a novel attention-based network

TCN output:

traffic prediction model, TSA-NTP, which effectively

captures the sp predictive accuracy. [7]

In this paper, we propose a novel attention-based network TCN output:

traffic prediction model, TSA-NTP, which effectively

captures the spatiotemporal features of complex network TCN output:

e In this paper, we propose a novel attention-based network

traffic prediction model, TSA-NTP, which effectively

captures the spatiotemporal features of complex network

construction of a graph Laplacian matrix to charact traffic prediction model, TSA-NTP, which effectively

captures the spatiotemporal features of complex network

environments. We provide a detailed discussion on the

construction of a graph Laplacian matrix to characteriz captures the spatiotemporal features of complex network

environments. We provide a detailed discussion on the

construction of a graph Laplacian matrix to characterize the

correlation between adjacent links within the n construction of a graph Laplacian matrix
correlation between adjacent links with
also generate a network traffic matrix i
validate the performance of the propos
experiments conducted under varying i
Our research not only Telation between adjacent links within the network. We compute the graph Laplaciation is a network traffic matrix using NSFNET and ideate a network traffice matrix using NSFNET and ideate the performance of the proposed So generate a network traffic matrix using NSFNET and

alidate the performance of the proposed method through

structure, and D is the adjacency

where A is the adjacence

were exact not only offers a new perspective on t validate the performance of the proposed method through
experiments conducted under varying network conditions. Structure, and D is the degree
or research not only offers a new perspective on traffic
provides theoretical

 $R^{N \times N}$, where N represents the number of nodes in the network. The historical network traffic data is represented as Experiments conducted under varying network conditions. structure, and D is the degree

Our research not only offers a new perspective on traffic explosion due to uncer degree

provides theoretical support for the future Our research not only offers a new perspective on traffic

prediction in complex network environments but also

provides theoretical support for the future deployment of AI-

provides the feature dimension technologies ai prediction in complex network environments but also

provides theoretical support for the future deployment of AI-

based predictive automation technologies aimed at managing

and complete the qury (Q),

2. **Methods**

In X^P based on the historical traffic X^T .
First, we apply a linear layer to the historical traffic data to rovides theoretical support for the future deployment of AI-
ased predictive automation technologies aimed at managing
nd simplifying network operations across all domains.

Dumber A in this study, we consider a network T_r and the state of the s Next, we apply the Laplacian

Simplifying network operations across all domains.

Hethods

In this study, we consider a network adjacency matrix $G \in \mathbb{R}^{T \times N \times T}$, where N represents the number of nodes in the

Next, w **2. Methods**
 2. Methods
 2. Methods
 3. Methods
 3. Methods
 4. In this study, we consider a network adjacency matrix $G \in \mathbb{W}_{0,KV}$ are the parameter of $W_{0,KV}$. The intermediate intervents the number of nod and simplitying network operations across all domains.

2. Methods

In this study, we consider a network adjacency matr
 $R^{N \times N}$, where N represents the number of nodes

network. The historical network traffic data is r resents the number of nodes in the

1 network traffic data is represented as

is the number of time steps in the

ic, and F is the feature dimension. The

2 prediction is to predict future traffic

cal traffic X^T .

ear

$$
H_L = Xw_1 + b_1
$$

$$
I_{i+1} = Re\,LU(Conv1D(H_i, W_i, d) + b_i + H_L)
$$

structure

eventuative weight matrix, and b_1 is the bias term.

ep enhances the model's capability to represent the

al features.

t, we employ a multi-layer Temporal Convolutional

k (TCN) to further model the process d b₁ is the bias term.
bility to represent the
emporal Convolutional
processed data. Each
, d) + b_i + H_L)
on, Conv1D represents
beration, W_i is the
ion rate. After passing
s, we obtain the final
capture the tempor **We del structure**
where W_1 is the weight matrix, and b_1 is the bias term.
is step enhances the model's capability to represent the
nporal features.
Next, we employ a multi-layer Temporal Convolutional
twork (TCN) t **model structure**
where W_1 is the weight matrix, and b_1 is the bias term.
This step enhances the model's capability to represent the
temporal features.
Next, we employ a multi-layer Temporal Convolutional
Network (T model structure
where W_1 is the weight matrix, and b_1 is the bias term.
This step enhances the model's capability to represent the
temporal features.
Next, we employ a multi-layer Temporal Convolutional
Network (TCN **model structure**

where W_1 is the weight matrix, and b_1 is the bias term.

This step enhances the model's capability to represent the

temporal features.

Next, we employ a multi-layer Temporal Convolutional

Netwo where W_1 is the weight matrix, and b_1
This step enhances the model's capability
temporal features.
Next, we employ a multi-layer Tempor
Network (TCN) to further model the proc
layer is represented as:
 $H_{i+1} = Re \, LU(Con$ is step enhances the model's capability to represent the
mporal features.
Next, we employ a multi-layer Temporal Convolutional
twork (TCN) to further model the processed data. Each
er is represented as:
 $H_{i+1} = Re LU(Conv1D(H_i,$ temporal features.

Next, we employ a multi-layer Temporal Convolutional

Network (TCN) to further model the processed data. Each

layer is represented as:
 $H_{i+1} = Re \, LU (Conv1D(H_i, W_i, d) + b_i + H_L)$

where ReLU is the activation f twork (TCN) to further model the processed data. Each

er is represented as:
 $H_{i+1} = Re \, LU (Conv1D(H_i, W_i, d) + b_i + H_L)$

where ReLU is the activation function, Conv1D represents

: one-dimensional convolution operation, W_i is th layer is represented as:
 $H_{i+1} = Re \, LU (Conv1D(H_i, W_i, d) + b_i + H_L)$

where ReLU is the activation function, Conv1D represents

the one-dimensional convolution operation, W_i is the

convolution kernel, and d is the dilation rate. where ReLU is the activation function, Conv1D represents
 \cdot one-dimensional convolution operation, W_i is the

nvolution kernel, and d is the dilation rate. After passing

ough multiple convolutional layers, we obtain where recently and any distribution profer the temperature intervals one-dimensional convolution operation. W₁ is the convolution kernel, and d is the dilation rate. After passing through multiple convolutional layers,

$$
H_{TCN} = H_r
$$

sequences.

$$
L = D - A
$$

d is the dilation rate. After passing

blutional layers, we obtain the final
 $H_{TCN} = H_n$

is designed to capture the temporal

orical traffic data, particularly in long

he network structure information, we

acian matrix explosive to unevenduced in the symmetrical convolution kernel, and d is the dilation rate. After passing through multiple convolutional layers, we obtain the final TCN output:
 $H_{TCN} = H_n$

The TCN module is designed to c EVALUATE: through multiple convolutional layers, we obtain the final
TCN output:
 $H_{TCN} = H_n$
The TCN module is designed to capture the temporal
dependencies in the historical traffic data, particularly in long
sequences.
T Exercise trains data, particularly in long

e the network structure information, we

placian matrix L, which is defined as:

L = D − A

acency matrix representing the network

the degree matrix. To prevent gradient

ven $H_{TCN} = H_n$
The TCN module is designed to capture the temporal
pendencies in the historical traffic data, particularly in long
quences.
To further leverage the network structure information, we
mpute the graph Laplacian ma The TCN module is designed to capture the temporal
dependencies in the historical traffic data, particularly in long
sequences.
To further leverage the network structure information, we
compute the graph Laplacian matrix dependencies in the historical traffic data, particularly i
sequences.
To further leverage the network structure information
compute the graph Laplacian matrix L, which is defined
 $L = D - A$
where A is the adjacency matrix r ere A is the adjacency matrix representing the network
ure, and D is the degree matrix. To prevent gradient
sion due to uneven degrees, we adopt the symmetrically
lized graph Laplacian matrix, expressed as:
 $\tilde{L} = I - D^{-\frac$ To further leverage the network structure information, we
mpute the graph Laplacian matrix L, which is defined as:
 $L = D - A$
where A is the adjacency matrix. To prevent gradient
ucture, and D is the degree matrix. To preven For the results of H_E D – A

Le D – A

Le D – A

where A is the adjacency matrix representing the network

structure, and D is the degree matrix. To prevent gradient

explosion due to uneven degrees, we adopt the symmetr

$$
\tilde{L} = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}
$$

 $\tilde{L} = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$
the Laplacian matrix to the feature matrix
tery (Q), key (K), and value (V) matrices
chanism:
 $U \left(H_{TCN}\tilde{L}w_{Q,K,V} + b_{Q,K,V} + H_{1\times1}\right)$
are the parameter matrices, and $H_{1\times1}^t$
to f H_L after $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$
an matrix to the feature matrix
ey (K), and value (V) matrices
 $w_{Q,K,V} + b_{Q,K,V} + H_{1\times1}$
arameter matrices, and $H_{1\times1}^t$
ter a 1x1 convolution:
 $\sum_{i=1}^C w_i \cdot H_i(t)$
computed as follows:
 V) = softmax The attention mechanism is computed as follows:
 $\tilde{L} = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$

Next, we apply the Laplacian matrix to the feature matrix

d compute the query (Q), key (K), and value (V) matrices

the attention mechanism:

$$
H_{Q,K,V} = ReLU \ (H_{TCN} \tilde{L} w_{Q,K,V} + b_{Q,K,V} + H_{1\times 1})
$$

$$
H_{1\times 1}^{t}=b+\sum_{i=1}^{C}w_{i}\cdot H_{i}(t)
$$

$$
H_{ATT} = \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V
$$

 $\tilde{L} = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$

t, we apply the Laplacian matrix to the feature matrix

mpute the query (Q), key (K), and value (V) matrices

attention mechanism:
 $I_{Q,K,V} = ReLU$ $(H_{TCN}\tilde{L}w_{Q,K,V} + b_{Q,K,V} + H_{1\times1})$

re $W_{Q,K,V}$ If dompute the query (Q), key (K), and value (V) matrices
the attention mechanism:
 $H_{Q,K,V} = ReLU$ ($H_{TCN} \tilde{L} w_{Q,K,V} + B_{Q,K,V} + H_{1\times1}$)
Where $W_{Q,K,V}$ are the parameter matrices, and $H_{1\times1}^t$
resents the result of H_L afte for the attention mechanism:
 $H_{Q,K,V} = ReLU$ $(H_{TCN} \tilde{L} w_{Q,K,V} + b_{Q,K,V} + H_{1 \times 1})$

Where $W_{Q,K,V}$ are the parameter matrices, and $H_{1 \times 1}^t$

represents the result of H_L after a 1x1 convolution:
 $H_{1 \times 1}^t = b + \sum_{i=1}^C w_i \$ $H_{1\times 1}^{t} = b + \sum_{i=1}^{C} w_i \cdot H_i(t)$
tion mechanism is computed as follows:
= Attention(Q, K, V) = softmax $\left(\frac{QK^{T}}{\sqrt{d_k}}\right)V$
ve perform residual connection and normalization
t of the attention mechanism and the TCN outp

ACK 2024 학술발표대회 논문집 (31권 2호)

After the above processing, the resulting feature vector is Based on the comparative anal

ssed through a Multi-Layer Perceptron (MLP) to produce TSANTP outperformed other mo

final predictio ACK 2024 학술발표대회 논문집 (31권 2호)

After the above processing, the resulting feature vector is

passed through a Multi-Layer Perceptron (MLP) to produce

the final prediction:
 $Y = MLP(H_{norm} + H_{MLP})$

where the MLP is defined as:
 After the above processing, the resulting features
passed through a Multi-Layer Perceptron (ML
the final prediction:
 $Y = MLP(H_{norm} + H_{MLP})$
where the MLP is defined as:
 $MLP(H) = W_f ReLU(HW_e + b_e) +$ After the above processing, the resulting feature vector is
seed through a Multi-Layer Perceptron (MLP) to produce TS
final prediction:
 $Y = MLP(H_{norm} + H_{MLP})$ int
where the MLP is defined as:
 $MLP(H) = W_f ReLU(HW_e + b_e) + b_f$ acl
Thus, we

$$
Y = MLP(H_{norm} + H_{MLP})
$$

$$
MLP(H) = W_f ReLU(HW_e + b_e) + b_f
$$

Y.

 $\frac{\text{ACK 2024}}{\text{Next: } \text{ACH}} = \frac{1}{2}$ $\text{ECH} = \frac{1}{2}$ $\text{Multi-Layer Perceptron (MLP) to produce} \qquad \text{TSAN'} \qquad \text{infeasible}$ $\text{INL} = \text{MLP}(H_{norm} + H_{MLP})$ $\text{P is defined as:} \qquad \text{infeasible} \\ \text{P} = \frac{W_f \text{ReLU}(HW_e + b_e) + b_f}{W_f \text{N}} = \frac{W_f \text{ReLU}(HW_e + b_e) + b_f}{W_f \text{NGRU}} \\ \text{in the final network traffic prediction result} \qquad \text{O.1$ = + + **Proposed through a Multi-Layer Perceptron (MLP) to produce**

TSANTP out

the final prediction:
 $Y = MLP(H_{norm} + H_{MLP})$

where the MLP is defined as:
 $MLP(H) = W_f ReLU(HW_e + b_e) + b_f$

Thus, we obtain the final network traffic prediction The mail prediction $Y = MLP(H_{norm} + H_{MLP})$

Where the MLP is defined as:
 $MLP(H) = W_f ReLU(HW_e + b_e) + b_f$ achieved

Thus, we obtain the final network traffic prediction result

The model, we introduce residual connections and LayerNorm
 To ensure the stability and generalization ability of the

methods of LayerNorm performed excellently v

alayers in the network, which effectively prevent gradient lower than other models

vanishing or explosion. Addition

model, we introduce residual connections and LayerNo
layers in the network, which effectively prevent gradio
vanishing or explosion. Additionally, we employ appropri-
weight initialization and regularization techniques (su Example to the method, we employ appropriate tower than other models. Additionally, we employ appropriate the network intensity was 15 cigarity may model to further improve the model's robustness.

Experimentation and Eva

weight initialization and regularization techniques (such as

weight initialization and regularization techniques (such as

demonstrating strong routing for

Dropout) to further improve the model's robustness.

This indic Dropout) to further improve the model's robustness.
 3. Experimentation and Evaluation

This indicates that my

This indicates that my

generalization ability and acc

The model is implemented using Pytorch-GPU 2.01 bas **3. Experimentation and Evaluation**

This indicates that my

generalization ability and according to the model is implemented using Pytorch-GPU 2.01 based

on Python 3.11 and is trained on a PC running Windows 11

Educati **3. Experimentation and Evaluation**

The model is implemented using Pytorch-GPU 2.01 based

on Python 3.11 and is trained on a PC running Windows 11

Education WSL, equipped with an Intel(R) Core(TM) i7-

IO700 CPU @ 2.90 The model is implemented using Pytorch-GPU 2.01 based

on Python 3.11 and is trained on a PC running Windows 11

Education WSL, equipped with an Intel(R) Core(TM) i7-

We compared our model wit

Idivide Gr practical

GPU, The model is implemented using Pytorch-GPU 2.01 based

on Python 3.11 and is trained on a PC running Windows 11

Education WSL, equipped with an Intel(R) Core(TM) i7-

We compared our model with

GPU, and 64 GB of memory. on Python 3.11 and is trained on a PC running Windows 11

Education WSL, equipped with an Intel(R) Core(TM) i7-

10700 CPU @ 2.90GHz, an Nvidia GeForce RTX 2060super

GPU, and 64 GB of memory.

We created a simulator usin chemes on NSFNET. We then generated
rix samples to reflect various data flows
NSFNET was operational from the 1980s
, connecting supercomputing centers and
Jnited States. It consisted of 14 nodes and
 V_e used MSE (Mean S 50,000 traffic matrix samples to reflect various data flows

both the network. NSFNET was operational from the 1980s

to the early 1990s, connecting supercomputing centers and

universities in the United States. It consis universities in the United States. It consisted of 14 nodes and TGCN

42 directed links. We used MSE (Mean Squared Error) as the similarly, length of capture evaluation metric, with the formula:

evaluation metric, with t

$$
MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
$$

 $\frac{\text{ACK 2024}}{\text{ACE 2024}}$

After the above processing, the resulting feature vector is

Sead on the comparative

Sead through a Multi-Layer Perceptron (MLP) to produce

TSANTP outperformed other

intensities and prediction ACK 2024 ^a \triangle \triangle 37 and 37 and 48 and through a Multi-Layer Perceptron (MLP) to produce

Similar production the comparative and frequency of the stability of the stability of the mall prediction is
 $Y = MLP(H_{norm} + H_{MLP})$
 Moreover \triangle and \triangle After the above processing, the resulting feature vector is

passed on the comparative

passed through a Multi-Layer Perceptron (MLP) to produce

the final prediction:
 $Y = MLP(H_{norm} + H_{MLP})$

where the MLP is defined as:
 $MLP(H)$ After the above processing, the resulting feature vector is

passed through a Multi-Layer Perceptron (MLP) to produce

the final prediction:
 $Y = MLP(H_{norm} + H_{MLP})$

where the MLP is defined as:
 $MLP(H) = W_fReLU(HW_e + b_e) + b_f$

Thus, we After the above processing, the resulting feature vector is

gassed on the comparative and

present in Amelyn a Multi-Layer Perceptron (MLP) to produce

the final prediction contraction $Y = MLP(H_{norm} + H_{MLP})$

where the MLP is $MLP(H) = W_f ReLU(HW_e + b_e) + b_f$

Thus, we obtain the final network traffic prediction result

Thus, we obtain the final network traffic prediction result

To ensure the stability and generalization ability of the intensity of 12 an Thus, we obtain the final network traffic prediction result

To ensure the stability and generalization ability of the

To ensure the stability and generalization ability of the

To ensure the stability and generalization For ensure the stability and generalization ability of the

model, we introduce residual connections and LayerNorm

model, we introduce residual connections and LayerNorm

layerS in the network, which effectively prevent $\frac{E - E}{\Delta}$ (31 <u>2 $\frac{2E}{\Delta}$)</u>
Based on the comparative analysis of the data in the table,
SANTP outperformed other models under different network
tensities and prediction lengths. When the network
tensity was 9 and th I \pm E \overline{a} (31 \overline{a} 2 \overline{c})

Based on the comparative analysis of the data in the table,

TSANTP outperformed other models under different network

intensities and prediction lengths. When the network

inte i $\pm \pm \pm \pm 2$ (31 ± \pideodotical dimensionally and prediction lengths. TSANTP outperformed other models under different network intensities and prediction lengths. When the network intensity was 9 and the prediction len inter and the comparative analysis of the data in the table,

TSANTP outperformed other models under different network

intensities and prediction lengths. When the network

intensity was 9 and the prediction length was 3, Based on the comparative analysis of the data in the table,
TSANTP outperformed other models under different network
intensities and prediction lengths. When the network
intensity was 9 and the prediction length was 3, my Based on the comparative analysis of the data in the table,
TSANTP outperformed other models under different network
intensities and prediction lengths. When the network
intensity was 9 and the prediction length was 3, my Based on the comparative analysis of the data in the table,

TSANTP outperformed other models under different network

intensities and prediction lengths. When the network

intensity was 9 and the prediction length was 3, **EEQ** (31 **H** 2 $\bar{\Sigma}$)

Based on the comparative analysis of the data in the table,

TSANTP outperformed other models under different network

intensities and prediction lengths. When the network

intensity was 9 and th **PERA (31 229)**

Based on the comparative analysis of the data in the table,

TSANTP outperformed other models under different network

intensities and prediction lengths. When the network

intensity was 9 and the predict Based on the comparative analysis of the data in the table,
TSANTP outperformed other models under different network
intensities and prediction lengths. When the network
intensity was 9 and the prediction length was 3, my Based on the comparative analysis of the data in the table,
TSANTP outperformed other models under different network
intensities and prediction lengths. When the network
intensity was 9 and the prediction length was 3, my Based on the comparative analysis of the data in the table,
TSANTP outperformed other models under different network
intensities and prediction lengths. When the network
intensity was 9 and the prediction length was 3, my TSANTP outperformed other models under different network
intensities and prediction lengths. When the network
intensity was 9 and the prediction length was 3, my model
achieved the lowest error of 0.101, whereas the errors intensities and prediction lengths. When the network
intensity was 9 and the prediction length was 3, my model
achieved the lowest error of 0.101, whereas the errors for
BIGRU, BiLSTM, STGCN, and TGCN were 0.111, 0.118,
0. intensity was 9 and the prediction length was 3, my model
achieved the lowest error of 0.101, whereas the errors for
BIGRU, BiLSTM, STGCN, and TGCN were 0.111, 0.118,
0.132, and 0.154, respectively. Similarly, under a netw achieved the lowest error of 0.101, whereas the errors for BIGRU, BiLSTM, STGCN, and TGCN were 0.111, 0.118, 0.132, and 0.154, respectively. Similarly, under a network intensity of 12 and a prediction length of 3, my model BIGRU, BiLSTM, STGCN, and TGCN were 0.111, 0.118,
0.132, and 0.154, respectively. Similarly, under a network
intensity of 12 and a prediction length of 3, my model again
performed excellently with an error of 0.088, signif 132, and 0.154, respectively. Similarly, under a network
tensity of 12 and a prediction length of 3, my model again
rformed excellently with an error of 0.088, significantly
wer than other models. Additionally, regardless intensity of 12 and a prediction length of 3, my model again
performed excellently with an error of 0.088, significantly
lower than other models. Additionally, regardless of whether
the network intensity was 15 or the pred performed excellently with an error of 0.088, significantly
lower than other models. Additionally, regardless of whether
the network intensity was 15 or the prediction length was 6,
my model consistently maintained the low lower than other models. Additionally, regardless of whether
the network intensity was 15 or the prediction length was 6,
my model consistently maintained the lowest error,
demonstrating strong robustness and predictive ca

In NSFNET. We then generated

es to reflect various data flows

Twas operational from the 1980s

ing supercomputing centers and

thes. It consisted of 14 nodes and

SE (Mean Squared Error) as the

states. It consisted of We created a simulator using OMNeT++ and applied 200

Terestiction steps of 3 and 6. As

Terent routing schemes on NSFNET. We then generated

0.000 traffic matrix samples to reflect various data flows

2.000 traffic matri different routing schemes on NSFNET. We then generated achieved the best performance 50,000 traffic matrix samples to reflect various data flows $\frac{1}{2}$ absolute Error). When the net within the network. NSFNET was opera **BIGRU:** BIGRU is a bidirectional GRU model that can be sequence data in both forward and backward are expected that sequence diversional from the 1980s inversities in the United States. It consisted of 14 nodes and targe to the early 1990s, connecting supercomputing centers and

universities in the United States. It consisted of 14 nodes and

42 directed links. We used MSE (Mean Squared Error) as the

evaluation metric, with the formula:
 Final and temporal dependencies in data. [8]

STGRU: STGCN: STGCN combines graph convolutional networks of the squared Error) as the signific the formula:

In the formula:
 $\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ Furthermo **EXECT:** STGCN combines graph convolutional networks of the series model in combine series and temporal dependences and temporal dependences of the predictional GRU model that can maintained the low capture temporal depen **EXECT:** STGCN combines graph convolutional networks

where $\frac{1}{n}$ and the special set of the prediction of the sequence.
 EXERU: BIGRU is a bidirectional GRU model that can maintained the lowest exapture temporal de MSE = $\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ Furthermore, regardless o
 TGCN: BIGRU is a bidirectional GRU model that can

maintained the lowest err

pture temporal dependency information from both the and predictive capabilit **BIGRU:** BIGRU is a bidirectional GRU model that can

capture temporal dependency information from both the and predictive capability

forward and backward directions of the sequence.
 BILSTM: BiLSTM is a bidirectional the network intensity was 15 or the prediction length was 6, my model consistently maintained the lowest error, demonstrating strong robustness and predictive capability. This indicates that my model possesses superior gen my model consistently maintained the lowest error,
demonstrating strong robustness and predictive capability.
This indicates that my model possesses superior
generalization ability and accuracy when handling tasks with
var demonstrating strong robustness and predictive capability.
This indicates that my model possesses superior
generalization ability and accuracy when handling tasks with
varying network intensities and prediction requirement This indicates that my model possesses superior
generalization ability and accuracy when handling tasks with
varying network intensities and prediction requirements,
making it suitable for practical applications.
We compar generalization ability and accuracy when handling tasks with
varying network intensities and prediction requirements,
making it suitable for practical applications.
We compared our model with four benchmark methods on
the varying network intensities and prediction requirements,
making it suitable for practical applications.
We compared our model with four benchmark methods on
the NSFNET dataset. Table 1 shows the results of prediction
perfo making it suitable for practical applications.
We compared our model with four benchmark methods on
the NSFNET dataset. Table 1 shows the results of prediction
performance at network intensities of 9, 12, and 15, with
pred We compared our model with four benchmark methods on
the NSFNET dataset. Table 1 shows the results of prediction
performance at network intensities of 9, 12, and 15, with
prediction steps of 3 and 6. As seen in Table 1, ou the NSFNET dataset. Table 1 shows the results of prediction
performance at network intensities of 9, 12, and 15, with
prediction steps of 3 and 6. As seen in Table 1, our TSANTP
achieved the best performance in terms of MA performance at network intensities of 9, 12, and 15, with
prediction steps of 3 and 6. As seen in Table 1, our TSANTP
achieved the best performance in terms of MAE (Mean
Absolute Error). When the network intensity was 9 an prediction steps of 3 and 6. As seen in Table 1, our TSANTP
achieved the best performance in terms of MAE (Mean
Absolute Error). When the network intensity was 9 and the
prediction length was 3, my model achieved the lowes achieved the best performance in terms of MAE (Mean
Absolute Error). When the network intensity was 9 and the
prediction length was 3, my model achieved the lowest error
of 0.101, whereas the errors for BIGRU, BiLSTM, STGC Absolute Error). When the network intensity was 9 and the prediction length was 3, my model achieved the lowest error of 0.101, whereas the errors for BIGRU, BiLSTM, STGCN, and TGCN were 0.111, 0.118, 0.132, and 0.154, res prediction length was 3, my model achieved the lowest error
of 0.101, whereas the errors for BIGRU, BiLSTM, STGCN,
and TGCN were 0.111, 0.118, 0.132, and 0.154, respectively.
Similarly, under a network intensity of 12 and of 0.101, whereas the errors for BIGRU, BiLSTM, STGCN, and TGCN were 0.111, 0.118, 0.132, and 0.154, respectively. Similarly, under a network intensity of 12 and a prediction length of 3, my model again performed excellent and TGCN were 0.111, 0.118, 0.132, and 0.154, respectively.
Similarly, under a network intensity of 12 and a prediction
length of 3, my model again performed excellently with an
error of 0.088, significantly lower than oth Similarly, under a network intensity of 12 and a prediction
length of 3, my model again performed excellently with an
error of 0.088, significantly lower than other models.
Furthermore, regardless of whether the network in length of 3, my model again performed excellently with an error of 0.088, significantly lower than other models.
Furthermore, regardless of whether the network intensity was 15 or the prediction length was 6, my model cons error of 0.088, significantly lower than other models.
Furthermore, regardless of whether the network intensity was
15 or the prediction length was 6, my model consistently
maintained the lowest error, demonstrating strong Furthermore, regardless of whether the network intensity was
15 or the prediction length was 6, my model consistently
maintained the lowest error, demonstrating strong robustness
and predictive capability. It can also be o 15 or the prediction length was 6, my mo
maintained the lowest error, demonstrating s
and predictive capability. It can also be
traditional time series analysis methods yield
while models considering both tempor
correlatio and predictive capability. It can also be
traditional time series analysis methods yielded
while models considering both temporal
correlations, such as STGCN and TGCN, were
This suggests that these methods have limit
model ile models considering both temporal and spatial
rrelations, such as STGCN and TGCN, were less effective.
is suggests that these methods have limited capacity in
odeling nonlinear and complex traffic data. Our TSANTP
ploys correlations, such as STGCN and TGCN, were less effective.
This suggests that these methods have limited capacity in
modeling nonlinear and complex traffic data. Our TSANTP
employs an attention mechanism and has outperform

This suggests that these methods have limited capacity in
modeling nonlinear and complex traffic data. Our TSANTP
employs an attention mechanism and has outperformed
previous state-of-the-art models, proving the advantages modeling nonlinear and complex traffic data. Our TSANTP
employs an attention mechanism and has outperformed
previous state-of-the-art models, proving the advantages of
our model in combining spatiotemporal features in netw employs an attention mechanism and has outperformed
previous state-of-the-art models, proving the advantages of
our model in combining spatiotemporal features in network
traffic prediction.
4. Conclusion
This study propo previous state-of-the-art models, proving the advantages of
our model in combining spatiotemporal features in network
traffic prediction.
4. Conclusion
This study proposes a network traffic prediction model
(TSA-NTP) bas our model in combining spatiotemporal features in network
traffic prediction.
4. Conclusion
This study proposes a network traffic prediction model
(TSA-NTP) based on the attention mechanism, which can
effectively capture traffic prediction.
 4. Conclusion

This study proposes a network traffic prediction model

(TSA-NTP) based on the attention mechanism, which can

effectively capture spatiotemporal characteristics in a

complex network accuracy and robustness of the prediction. Experimental

ACK 2024 학술발표대회 논문집 (31권 2호)
has the lowest error, showing superior generalization ability
and prediction performance. In contrast, traditional time
series analysis methods and models that only consider
spatiotemporal cor ACK 2024 학술발표대회 논문집 (31권 2호)
has the lowest error, showing superior generalization ability
and prediction performance. In contrast, traditional time
series analysis methods and models that only consider
spatiotemporal cor ACK 2024 학술발표대회 논문집 (31권 2호)
has the lowest error, showing superior generalization ability
and prediction performance. In contrast, traditional time
series analysis methods and models that only consider
spatiotemporal corr ACK 2024 학술발표대회 논문집 (31권 2호)
has the lowest error, showing superior generalization ability
and prediction performance. In contrast, traditional time
series analysis methods and models that only consider
spatiotemporal cor $\frac{\text{ACK 2024}}{\text{Next 2024}} \times \frac{\text{Text 2024}}{\text{Text 2024}} \times \frac{\text{Text 2024}}{\text{Text 2024}}$

has the lowest error, showing superior generalization ability

and prediction performance. In contrast, traditional time

series analysis methods a ACK 2024 학술발표대회 논문집 (31권 2호)
has the lowest error, showing superior generalization ability
and prediction performance. In contrast, traditional time
series analysis methods and models that only consider
spatiotemporal cor provides theoretical support for artificial intelligence-based $ACK 2024$ $\triangleq \triangleq \triangleq \triangleq \triangleq H \triangleq (31 \triangleq 2 \triangleq)$
has the lowest error, showing superior generalization ability
and prediction performance. In contrast, traditional time
series analysis methods and models that only consider
spa $\frac{\text{ACK 2024}}{\text{ASE}}$ has the lowest error, showing superior generalization ability
and prediction performance. In contrast, traditional time
series analysis methods and models that only consider
spatiotemporal correlation management. d prediction performance. In contrast, traditional time
ricis analysis methods and models that only consider
atiotemporal correlations do not perform as well as TSA-
TP in processing nonlinear and complex traffic data. Thi series analysis methods and models that only consider
spatiotemporal correlations do not perform as well as TSA-
NTP in processing nonlinear and complex traffic data. This
shows that the model proposed in this study not on spatiotemporal correlations do not perform as well as TSA-
NTP in processing nonlinear and complex traffic data. This
shows that the model proposed in this study not only
provides theoretical support for artificial intelli

Acknowledgement

NTP in processing nonlinear and complex traffic data. This
shows that the model proposed in this study not only
provides theoretical support for artificial intelligence-based
prediction automation technology, but also has shows that the model proposed in this study not only
provides theoretical support for artificial intelligence-based
prediction automation technology, but also has the potential
to be widely used in actual network operation provides theoretical support for artificial intelligence-based
prediction automation technology, but also has the potential
to be widely used in actual network operation and
management.
This work was supported by Innovativ prediction automation technology, but also has the potential
to be widely used in actual network operation and
management.
This work was supported by Innovative Human Resource
Development for Local Intellectualization prog to be widely used in actual network operation and
management.

This work was supported by Innovative Human Resource

Development for Local Intellectualization program through

the Institute of Information & Communications **Acknowledgement**
 Acknowledgement

This work was supported by Innovative Human Resource

Development for Local Intellectualization program through

the Institute of Information & Communications Technology

Planning & Ev Acknowledgement

This work was supported by Innovative Human Resource

Development for Local Intellectualization program through

the Institute of Information & Communications Technology

Planning & Evaluation(IITP) grant government(MSIT). **Framing & Evaluation(IITP)** grant funded by the Korea
government(MSIT)(IITP-2024-RS-2022-00156287, 50%).
This work was supported by Institute of Information &
communications Technology Planning & Evaluation (IITP)
under Framing C Evantaton(ITTP-2024-RS-2022-00156287, 50%).

This work was supported by Institute of Information &

communications Technology Planning & Evaluation (IITP)

under the Artificial Intelligence Convergence Innovatio Boven This work was supported by Institute of Information 8

communications Technology Planning & Evaluation (IITP

under the Artificial Intelligence Convergence Innovation

Human Resources Development (IITP-2023-RS-2023
 Inis work was supported by Institute of Information &
communications Technology Planning & Evaluation (IITP)
under the Artificial Intelligence Convergence Innovation
100256629, 50%) grant finded by the Korea
government(MS communications 1echnology Planning & Evaluation (IITP)

under the Artificial Intelligence Convergence Innovation

Human Resources Development (IITP-2023-RS-2023-

00256629, 50%) grant funded by the Korea

government(MSIT) under the Artificial Intelligence Convergence Innovation
Human Resources Development (IITP-2023-RS-2023-
00256629, 50%) grant funded by the Korea
government(MSIT).
 $\Delta \overline{\Delta} = \overline{\Delta}$
[1] Joshi M, Hadi T H. A review of netw

참고문헌

Human Resources Development (IITP-2023-RS-2023-

00256629, 50%) grant funded by the Korea

government(MSIT).
 $\Delta^2 \mathbf{L} \equiv \mathbf{\hat{e}} \mathbf{\hat{f}}$

[1] Joshi M, Hadi T H. A review of network traffic

analysis and prediction tec sports and Management (MSIT).
 $\Delta L = 2$

[1] Joshi M, Hadi T H. A review of network traffic

analysis and prediction techniques[J]. arXiv preprint

arXiv:1507.05722, 2015.

[2] Li L, Kim K. GTT-NTP: A Graph Convolutional
 and The Soliton School S $\begin{array}{ll}\n & \text{A} \rightarrow \mathbb{R} \oplus \mathbb{R} \longrightarrow \mathbb{$ **EDERT ASSET ASSET ASSET ASSET ASSET ASSET AND A THE STATIST AND ARRIVITION AND NEW AND NUMBER (21) Li L, Kim K. GTT-NTP: A Graph Convolutional Network-Based Metwork Traffic Prediction model [C]/NOMS 2024-2024 IEEE Networ aggregated**
 aggregates
 aggregated in the service of network traffic

analysis and prediction techniques[J]. arXiv preprint

arXiv:1507.05722, 2015.

[2] Li L, Kim K. GTT-NTP: A Graph Convolutional

Networks-Based M [1] Joshi M, Hadi T H. A review of network traffic
analysis and prediction techniques[J]. arXiv preprint
arXiv:1507.05722, 2015.
[2] Li L, Kim K. GTT-NTP: A Graph Convolutional
Networks-Based Network Traffic Prediction
mo [1] Joshi M, Hadi I H. A review of network traftic
analysis and prediction techniques[J]. arXiv preprint
arXiv:1507.05722, 2015.
[2] Li L, Kim K. GTT-NTP: A Graph Convolutional
Networks-Based Network Traffic Prediction
mo arXiv:1507.05722, 2015.

[2] Li L, Kim K. GTT-NTP: A Graph Convolutional

Networks-Based Network Traffic Prediction

model[C]//NOMS 2024-2024 IEEE Network Operations

and Management Symposium. IEEE, 2024: 1-7.

[3]Mukherje [2] Li L, Kim K. GTT-NTP: A Graph Convolutional
Networks-Based Network Traffic Prediction
model[C]//NOMS 2024-2024 IEEE Network Operations
and Management Symposium. IEEE, 2024: 1-7.
13] Mukherjee S, Ray R, Samana R, et al

Networks-Based Network Traffic Prediction
model[C]//NOMS 2024-2024 IEEE Network Operations
and Management Symposium. IEEE, 2024: 1-7.
[3]Mukherjee S, Ray R, Samanta R, et al. Nonlinearity and
chaos in wireless network traf model[C]//NOMS 2024-2024 IEEE Network Operations
and Management Symposium. IEEE, 2024: 1-7.
[3]Mukherjee S, Ray R, Samanta R, et al. Nonlinearity and
chaos in wireless network traffic[J]. Chaos, Solitons &
Fractals, 2017, and Management Symposium. IEEE, 2024: 1-7.

[3]Mukherjee S, Ray R, Samanta R, et al. Nonlinearity and

chaos in wireless network traffic[J]. Chaos, Solitons &

Fractals, 2017, 96: 23-29.

[4]Lazaris A, Prasanna V K. Deep l [3]Mukherjee S, Ray R, Samanta R, et al. Nonlinearity and
chaos in wireless network traffic[J]. Chaos, Solitons &
Fractals, 2017, 96: 23-29.
[4]Lazaris A, Prasanna V K. Deep learning models for
aggregated network traffic Fractals, 2017, 96: 23-29.

[4] Lazaris A, Prasanna V K. Deep learning models for

aggregated network traffic prediction [C]//2019 15th

International Conference on Network and Service

Management (CNSM). IEEE, 2019: 1-5.
 [4]Lazaris A, Prasanna V K. Deep learning models for
aggregated network traffic prediction[C]//2019 15th
International Conference on Network and Service
Management (CNSM). IEEE, 2019: 1-5.
[5]Yeom S, Choi C, Kolekar S S, e

regregated network traffic prediction[C]//2019 15th
International Conference on Network and Service
Management (CNSM). IEEE, 2019: 1-5.
[5]Yeom S, Choi C, Kolekar S S, et al. Graph convolutional
network based link state pr Experimental Conference on Network and Service
Management (CNSM). IEEE, 2019: 1-5.
[5]Yeom S, Choi C, Kolekar S S, et al. Graph convolutional
network based link state predcition[C]//2021 22nd Asia-
Pacific Network Operatio Management (CNSM). IEEE, 2019: 1-5.

[5]Yeom S, Choi C, Kolekar S S, et al. Graph convolutional

network based link state prediction[C]//2021 22nd Asia-

Pacific Network Operations and Management Symposium

(APNOMS). IEEE, Management (CNSM). IEEE, 2019: 1-5.
[5]Yeom S, Choi C, Kolekar S S, et al. Graph convolutional
network based link state prediction[C]//2021 22nd Asia-
Pacific Network Operations and Management Symposium
(APNOMS). IEEE, 202

81.

[5]Yeom S, Choi C, Kolekar S S, et al. Graph convolutional
network based link state prediction[C]//2021 22nd Asia-
Pacific Network Operations and Management Symposium
(APNOMS). IEEE, 2021: 246-249.
[6]Zhou J, Cui G, Hu S, network based link state prediction[C]//2021 22nd Asia-
Pacific Network Operations and Management Symposium
(APNOMS). IEEE, 2021: 246-249.
[6]Zhou J, Cui G, Hu S, et al. Graph neural networks: A
review of methods and appli Pacific Network Operations and Management Sym
(APNOMS). IEEE, 2021: 246-249.
[6]Zhou J, Cui G, Hu S, et al. Graph neural netw
review of methods and applications[J]. AI open, 2020
81.
[7]Vaswani A. Attention is all you need