Journal of Systems Engineering and Electronics
Vol. 35, No. 4, August 2024, pp.855 – 864

Early warning of core network capacity in space-terrestrial integrated networks

HAN Sai[1,*], LI Ao[1], ZHANG Dongyue[1], ZHU Bin[1], WANG Zelin[1], WANG Guangquan[1], MIAO Jie[2], and MA Hongbing[2]

1. China Unicom Research Institute, Beijing 100044, China;
2. China United Network Communications Group Company Limited, Beijing 100033, China

Abstract: With the rapid development of low-orbit satellite communication networks both domestically and internationally, space-terrestrial integrated networks will become the future development trend. For space and terrestrial networks with limited resources, the utilization efficiency of the entire space-terrestrial integrated networks resources can be affected by the core network indirectly. In order to improve the response efficiency of core networks expansion construction, early warning of the core network elements capacity is necessary. Based on the integrated architecture of space and terrestrial network, multidimensional factors are considered in this paper, including the number of terminals, login users, and the rules of users' migration during holidays. Using artifical intelligence (AI) technologies, the registered users of the access and mobility management function (AMF), authorization users of the unified data management (UDM), protocol data unit (PDU) sessions of session management function (SMF) are predicted in combination with the number of login users, the number of terminals. Therefore, the core network elements capacity can be predicted in advance. The proposed method is proven to be effective based on the data from real network.

Keywords: space-terrestrial integrated networks, core network, element capacity, artificial intelligent (AI).

DOI: 10.23919/JSEE.2024.000072

1. Introduction

With the rapid development of low-orbit satellite [1] communications, space networks, and terrestrial networks are gradually forming two independent information communication networks. Although space networks have wide coverage areas [2] and robust disaster tolerance, network performance can be greatly affected by environmental conditions, terrestrial weather, and buildings. Currently, the system capacities [3] and speeds are far lower than those of terrestrial 5th generation mobile communication technology (5G) networks. Although terrestrial 5G networks have high performance, big capacity, and excellent service quality, they cannot be applied in remote areas, such as the scenes of oceans, mountains/suburbs with problems of business continuity and ubiquitous service. At present, there is a trend of integrating satellite and 5G networks. As a supplementary extension, satellite communications enhances the capabilities of 5G networks, and space networks extend the spatial dimension of the terrestrial networks, enabling global coverage communication with high reliability and high security [4].

The integration of space and terrestrial networks plays an important role in the research areas of enhanced 5G and 6th generation mobile communication technology (6G) scenarios [5,6]. In the space-terrestrial integrated networks (STIN), the advantages of both networks are combined and an effective solution for "always-online" communication is provided.

From the perspective of long-term development of integrated networks, space networks mainly include space access networks and space core networks, while the terrestrial networks mainly include terrestrial access networks and terrestrial core networks, respectively [6]. In both space networks and terrestrial networks, the core networks are the core of network control nodes and have a great effect on controlling user registration, session management, mobility management, traffic control and so on. Therefore, the utilization efficiency of STIN resources can be affected by the core networks indirectly. For space networks or terrestrial networks, it is necessary to improve the response efficiency of expansion and construction for core networks.

Recently, with the development of digital transformation, the artificial intelligence (AI) technologies have been widely used in the communication networks [7], covering the whole processes of network planning, con-

Manuscript received April 10, 2023.
*Corresponding author.
This work was supported by the National Key Research Plan (2021YFB2900602).

struction, maintenance, optimization, and operating, for example, fault prediction [7−10], performance prediction [11], energy conservation of base station [12], business model prediction [13], load balancing [14], network anomaly detection, and capacity prediction [15,16].

AI technologies have been applied to core networks as well. Long short-term memory (LSTM), gradient boosting decision tree (GBDT), autoregressive integrated moving average (ARIMA) model, and Prophet algorithms [17–21] have been used to predict the central processing unit (CPU) and memory usage of access and mobility management function (AMF), session management function (SMF), and user plane function (UPF) network elements in core networks [15,18,19,21]. However, these research focus on CPU and memory usage. There is no research on predicting the capacities of core network elements (AMF, SMF, unified data management (UDM)) yet, especially combined with the number of login users and the number of terminals.

Considering multi-dimensional factors such as the number of user terminals, the number of login users, and the rules of user migration during holidays, the capacities of 5G core network elements (AMF, SMF, UDM) are predicted in this paper. Network demand capacity in the future can be predicted in advance, which contributes to the rapid development of 5G and voice over long-term evolution (VoLTE) users. Therefore, the health of network load can be monitored in real time with dynamic thresholds, and the capacity expansion cost can be reduced. Furthermore, the utilization rate of network resources, as well as the response efficiency of capacity expansion construction can be effectively improved.

This paper is organized as follows. An overview of STIN is depicted in Section 2, including network architecture and evolution routes of STIN. The method of core network capacity warning is proposed in Section 3, consisting technical scheme and algorithm. Implementation and results are described in Section 4. Finally conclusions are drawn in Section 5.

2. An overview of STIN

2.1 Network architecture

With 5G and space networks being incorporated into the “new infrastructure” category in China, STIN are the future trend of new information infrastructure. More and more research focus on improving the coverage of STIN and offering “always-online” capabilities [22–24]. This can be achieved by utilizing information technology resources and respective advantages efficiently, and solving the problems of high dynamic network topology, mixed business differentiation, and difficulty in platform heterogeneous interoperability.

The overall architecture of STIN meets the requirements of ubiquitous multi-level access in the future. The STIN architecture is shown in Fig. 1.

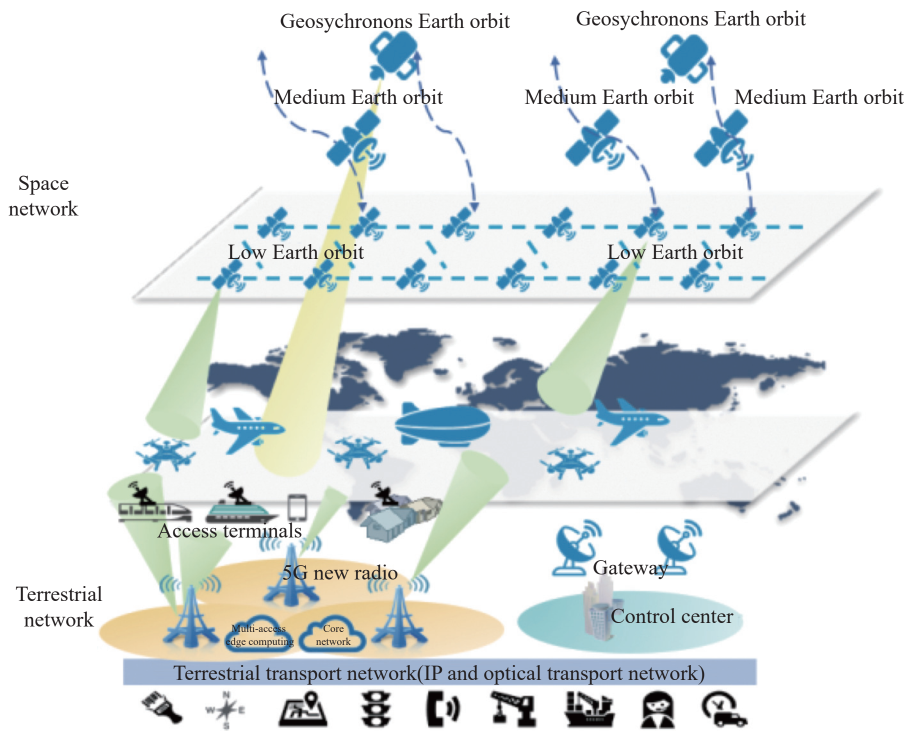

Fig. 1 STIN architecture

STIN architecture consists of the following systems [25–29]:

(i) Space system with communication, navigation, and remote sense

i) Remote sensing satellite: space information;

ii) Navigation satellite: dynamic information;

iii) Communication satellite: communication information;

iv) Satellite plus highaltitude platform station (HAPS): constellations, high-altitude airships, drone, aircraft, and other platforms.

(ii) Terrestrial network (transmission) system

i) Interchange (access / fusion) gateway: space remote communication station and terrestrial gateway;

ii) Terrestrial stations, transmission, and core networks.

2.2 Evolution routes of STIN

From the perspective of fixed, mobile, and space-integrated access, the third generation partner project (3GPP) protocols are mainly used in terrestrial networks, including fixed and mobile access. The non-3GPP access methods are also supported at the same time. In contrast, non-3GPP protocols such as digital video broadcast (DVB) are initially adopted in space networks. The two systems are developed independently initially and interconnected from the terminal to the access networks. From the perspective of long-term development, the evolution of integration can be divided into three stages: business integration, system integration, and deep integration. These stages are shown in Fig. 2 [24,30–34].

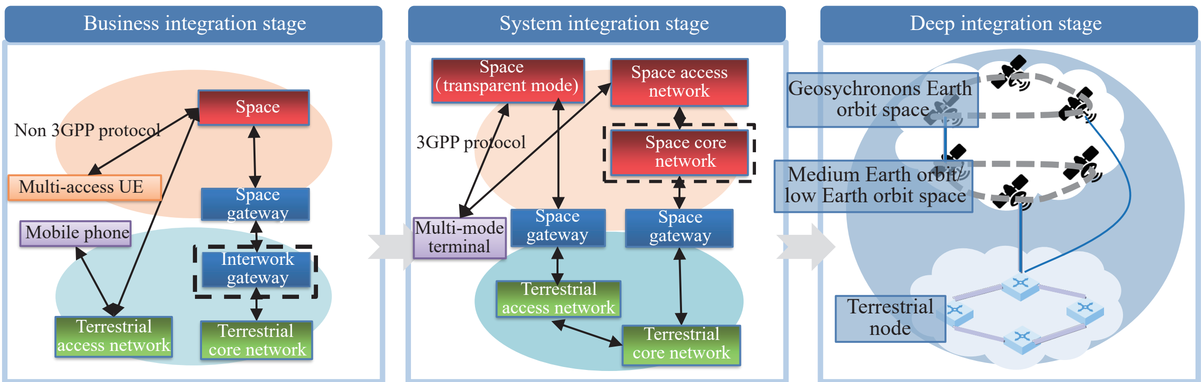

Fig. 2 Three stages of integration and evolution

Stage 1 Business integration stage

Integration mode: multi-mode terminals, network interworking, and private satellite protocol;

Development strategy: as an emergency supplement to real networks, rely on narrowband communication capabilities of satellite in orbit, in order to solve the emergency communication requirements such as short messages and voice in remote areas without cellular scenarios [30–34].

Stage 2 System integration stage

Integration mode: some of network elements is integrated with satellite, and the 3GPP standard communication system is adopted in STIN.

Development strategy: the coverage capability of cellular networks can be enhanced by STIN, providing users with seamless coverage voice and wide/narrow band mobile communication services over the entire domains [30–34].

Stage 3 Deep integration stage

Integration mode: toward 6G standard, network elements are on-board based on demand and deployed in a distributed manner.

Development strategy: based on beyond 5G (B5G) system integration, the network elements are deployed on demand and AI on-board to provide terrestrial cellular users with high reliability and high-quality intelligent networks. Considering the limited resource and capacity of space nodes, as well as the topological time-varying nature of the satellite constellation of non-geostationary satellite orbit (NGSO), it is necessary to efficiently utilize the physical resources of the STIN and predict the capacity and bandwidth usage of corresponding network elements in advance. During the initial adaptation, intermediate use, and later release of STIN network resources, AI algorithms are introduced to predict network element resources and assist in the intelligent management of network elements. In order to predict capacity, first of all, the resource pool of the STIN should be abstracted to form differential space and terrestrial resource supply capabilities; secondly, the network topology and the capacity of space nodes' resources are imported. Finally, AI algorithms are introduced to predict resource capacity requirements [30–34].

Due to the reason that the lack of satellite data, the network capacity management of terrestrial network is studied in this paper.

3. Core network capacity warning

The core network is one of the important networks in the STIN. In order to bear the significant increase in data traffic and ensure the expansion needs of data-intensive services, the number of 5G core network devices, which play an important role in global resource management and data service carrying, will be sharply increased. The implementation cycle of expansion takes a long time. It is often difficult to respond to market expansion needs timely. The capacity of core network elements and expansion plans should be predicted in advance to ensure user perception and alleviate the pressure of future traffic increase.

3.1 Technical scheme

In order to solve the urgent demand of network for the rapid development, a capacity early warning and scheduling module is developed on the network functions virtualization orchestration (NFVO) system, which can give early warning, once the network element capacity is beyond the threshold. The technical scheme architecture is shown in Fig. 3.

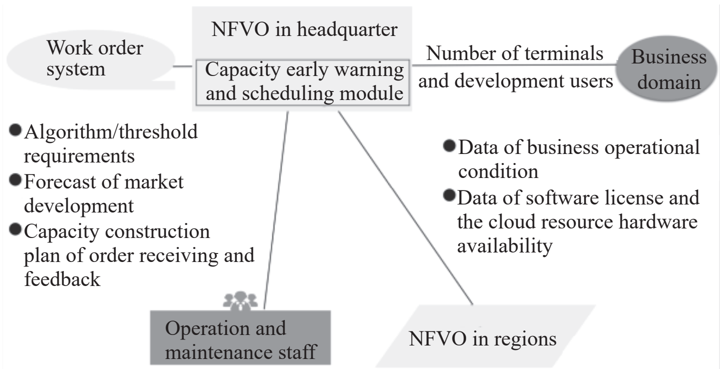

Fig. 3 Technical scheme of capacity warning and scheduling module

A two-level architecture with headquarter and regions is adopted in this system. The NFVO in headquarter is the main management platform of the core network, providing management and maintenance abilities for the whole company's 5G core and other core network sub-domains. The data of business operational condition, software license, and the cloud resource hardware availability is reported per hour by the NFVO in regions, to the NFVO in headquarter. The algorithm/threshold requirements are reported by the operation and maintenance staff to the NFVO in headquarter. The forecast of market development, the capacity construction plan of order receiving and feedback can also be imported later. In addition, the number of terminals in business domain and development users can also be provided to the NFVO headquarter. Based on these data, the network element capacity can be predicted by the capacity early warning and scheduling module comprehensively.

3.2 Algorithm

3.2.1 Prophet

Prophet is an algorithm designed to predict single-variable time series data sets. In the Prophet model, the prior knowledge of multiple variables is used to predict the time series of a single variable and with nonlinear characteristics. Comparted to using the LSTM algorithm, the result of using the Prophet algorithm is quite similar. However, the training cost and the using difficulty is lower. The fitting speed of the Prophet algorithm is very fast, and various different combinations can be tried in a limited time. The overall performance of using Prophet is relatively better, therefore, the Prophet is taken in this paper. Unlike traditional time series prediction methods, a new solution to transform the prediction problem into a fitting problem is proposed in the Prophet algorithm. Based on the decomposition of the time series as an additive model, the trend term, cycle term, and holiday term are decoupled to solve the time series prediction problem [35,36].

The Prophet model can be decomposed as

$$y(t)=g(t)+s(t)+h(t)+\varepsilon_t \tag{1}$$

where $g(t)$ is the trend function with non-periodic changes, $s(t)$ describes the periodic function with periodic changes, i.e., seasonality; $h(t)$ represents the effects of holidays with potential irregular schedules, and ε_t are the unique changes not fitted by the model. The detailed information can be found in [35].

In this paper, a number of three network element capacity, the number of login users (NU), the number of terminals (NT) are studied. For each network element capacity from one province, the length of sliding window from history input data is defined with inlen, the length of interval between the input data and output data is defined with space, the length of the output date is noted with outlen, D is the data including network element capacity, NU, and TU. In that way, data $D[i{:}i+\text{inlen}-1]$ used to predict $D[i+\text{inlen}+\text{space}{:}i+\text{inlen}+\text{space}+\text{outlen}-1]$, with i being the starting date index. The prophet algorithm is implemented by the following steps.

Step 1 Data pre-processing

The average values of the adjacent data are interpolated once there is empty data. After data preprocessing, the data is normalized respectively and divided into the training and test sets. The data format after preprocessing is set as D:[date, network element capacity, NU, NT] with four dimensions.

Step 2 Parameter configuration

The model hyper-parameters is configured as follows:
yearly_seasonality=40,
growth=‘linear’,
changepoints=None,
*n*_changepoints=25,
changepoint_range=0.8,
yearly_seasonality=40,
weekly_seasonality=‘auto’,
daily_seasonality=‘auto’,
seasonality_mode=‘additive’,
seasonality_prior_scale=10.0,
holidays_prior_scale=10.0,
changepoint_prior_scale=0.05,
mcmc_samples=0,
interval_width=0.80,
uncertainty_samples=1 000,
stan_backend=None.

3.2.2 DeepState

In the analysis of time series, some classic methods, such as ARIMA and Holt Winters, can be rebuilt as state space models. The state space model can be modeled each time series separately and similar patterns between sequences cannot be utilized, therefore, it is ineffective for limited historical data.

The state space model is combined with deep learning in DeepState model. The recurrent neural network is used to map the features into the parameters of the state space model, which can predict the probability distribution of the sequence values at each time index. Similar patterns from a large number of series and features are learned, and the model is therefore explainable [37,38].

Moreover, there is stronger prior knowledge in Deepstate, therefore, less training data should be required. In this paper, the historical data is only about one year, there are also some missing values. A robust algorithm should be chosen. Various algorithms haven been tried in this paper, and Deepstate is found to have a better performance than DeepAR and other algorithms, therefore, DeepState is compared with Prophet in this paper.

A state transition and an observation model is included in the state space model. The rule of hidden state changing over time is described in a state transition, with the transition probability $p(\boldsymbol{l}_t|\boldsymbol{l}_{t-1})$, and latent state $\boldsymbol{l}_t \in \mathbf{R}^L$, where t is the time index in the training range of $\{1,2,\cdots,T\}$ [37].

For the observation model, the conditional probability of the observations under a given latent state $p(z_t|\boldsymbol{l}_t)$ is defined, with z_t being the Gaussian observation model.

The process of the DeepState model is performed as follows.

Firstly, the recurrent neural network (RNN) model is used to calculate the network output:

$$\boldsymbol{h}_t = \mathrm{RNN}(\boldsymbol{h}_{t-1}, \boldsymbol{x}_t) \tag{2}$$

where $\boldsymbol{x}_t$ is the feature and assumed as a covariate vector, $\boldsymbol{h}_{t-1}$ is the previous network output.

Then the value $\boldsymbol{h}_t$ is used to calculate the parameters of the state space model:

$$\boldsymbol{\theta}_t = \varphi(\boldsymbol{h}_t). \tag{3}$$

Finally, the likelihood is estimate by

$$\mathbf{pss} = (z_{1:T}|\boldsymbol{\theta}_{1:T}) \tag{4}$$

where **pss** is calculated marginal probability of a given observation value $z_{1:T}$ using parameter $\boldsymbol{\theta}_{1:T}$, $z_{1:T}$ is target time series (observations value) from time 1 to T, $\boldsymbol{\theta}_{1:T}$ is the parameter of state space model from 1 to T [37].

The future data can be predicted by maximizing the log-likelihood learning network parameters. The detailed information can be found in [37].

In this paper, the DeepState model is implemented by the following steps:

Step 1 Data preprocessing

The same as Step 1 in Subsection 3.2.1.

Step 2 Parameter configuration

The model hyper-parameters are configured as follows:
freq=“*D*”,
prediction_length=30,
epochs=20,
cell_type= “lstm”
num_batches_per_epoch=50,
use_feat_dynamic_real=True.

3.3 Model evaluation

In order to evaluate the performance of the prediction, two measurements, one is coefficient of determination R^2, and the other is the mean square error (MSE), are adopted in this paper.

3.3.1 Coefficient of determination

The performance of a predictor can be indicated by the coefficient of determination R^2. The correlation between the observed values and the predicted values of the regression model can be illustrated by the percentage of effective variance. The mean value is often used to be the benchmark error. R^2 [39] is calculated as follows:

$$R^2 = 1 - \frac{\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{n}(y_j - \bar{y}_j)^2} \tag{5}$$

where y_j represents the actual observation value, $\bar{y}_j$ is the average value, $\hat{y}_j$ denotes the predicted value, and j is the time index .

If R^2=0, the numerator equals the denominator, and each predicted sample value equals the mean value. If R^2=1, the predicted value and the true value in the sample are entirely same without any error [39].

3.3.2 MSE

MSE is based on the square of the difference between the actual and predicted values. It is often used as the loss function of linear regression. The smaller the value of MSE is, the better the accuracy of the prediction model to describe the experimental data is. MSE [40] is computed by

$$\mathrm{MSE} = \frac{1}{m}\sum_{j=1}^{m}(y_j - \hat{y}_j)^2 \tag{6}$$

where m represents the total amount of data.

3.4 Correlation analysis

Before performing the aforementioned algorithm, the cross-correlation function (CCF) was used in this paper to analyze the correlation between each feature and each network elements capacity, aiming to select the features that significantly affect the capacity of each network element. CCF [41] can be obtained by

$$\mathrm{ccf}_n = \frac{\sum_{t=1}^{N}\left(X_t - \bar{X}\right)\left(Y_{t+n} - \bar{Y}\right)}{\sum_{t=1}^{N}\left(X_t - \bar{X}\right)^2\left(Y_{t+n} - \bar{Y}\right)^2} \tag{7}$$

where X_t represents the value of time series $\boldsymbol{X}$ at time t, Y_{t+n} denotes the value of time series $\boldsymbol{Y}$ at time $t+n$, n is the time difference, N implies the number of data, $\bar{X}$ and $\bar{Y}$ are the average value of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. The ccf indicates the degree of correlation between $\boldsymbol{X}$ and $\boldsymbol{Y}$ [41].

The closer the absolute value of the CCF result is to 1, the stronger the correlation can be discovered. Moreover, it is known that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are negatively correlated if the result of CCF is negative [41].

4. Implementation and results

Data from September 01, 2021 to Novermber 06, 2022 in one province taken from China Unicom real network is used in this paper, with a number of 414 in total. The data is collected once per day, and includes three network elements: AMF, UDM, and SMF. The network elements capacity refer to the number of registered users of AMF, the authorization users of UDM, and protocol data unit (PDU) sessions of SMF, respectively. All data is arranged in a chronological order.

Applying the data format from Subsection 3.2.1 and Subsection 3.2.2, a number of 384 data (from September 01, 2021 to October 07, 2022) is used for training, and a number of 30 data (from October 08, 2022 to November 06, 2022) is used to be predicted.

4.1 Results of correlation analysis

The correlation analysis between each network element capacity and each feature is based on the data sequences of AMF, UDM, and SMF network elements capacity and features from September 01, 2021 to April 17, 2022. The features include holiday, workday, weekend, NU, and NT. The correlation analysis results are shown in Table 1.

Table 1 Results of correlation analysis

Feature	AMF/Ten thousand users	UDM/Ten thousand users	SMF/ Ten thousands sessions
NU/Ten thousand users	0.876026	0.985335	0.847536
NT/Ten thousand users	0.871045	0.968826	0.833428
Holiday	−0.264972	−0.027671	−0.299663
Workday	−0.143109	−0.008093	−0.180952
Weekend	−0.00803	0.000481	−0.028388

Aiming to describe the correlations between each network element capacity and each feature more intuitively, the scatter diagrams are plotted. Some of the scatter diagrams are shown in Fig. 4 and Fig. 5. The closer the scatter plot to a straight line is, the higher the correlation exists. With an example of holidays and AMF shown in Fig. 4, it can be inferred that there is no correlation between holidays and the value of AMF. The results of UDM and NU in depicted in Fig. 5, it can be clearly seen that the UDM is strongly positively correlated with login users.

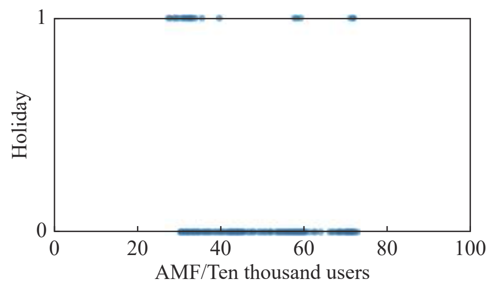

Fig. 4 Scatter diagram of AMF and holiday

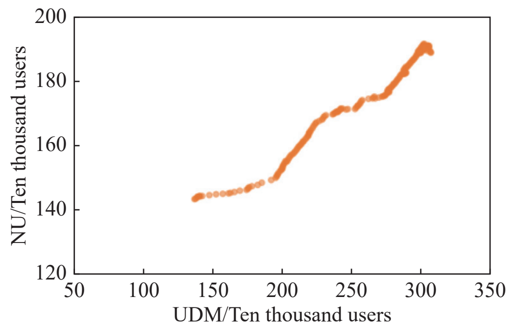

Fig. 5 Scatter diagram of UDM and NU

It can be concluded that there is strong and positive correlations between the network element capacity of AMF, UDM, and SMF with NT and NU, and there is no correlation between holiday, workday, or weekend. Therefore, NT and NU are added to the algorithms for predicting the network element capacity of AMF, UDM, and SMF.

4.2 Results of algorithm

4.2.1 AMF

The true value, the predicted value using Prophet algorithm, and the predicted value using DeepState algorithm for the registered users of AMF are shown in Fig. 6. The results of R^2 and MSE to evaluate the accuracy are depicted in Table 2.

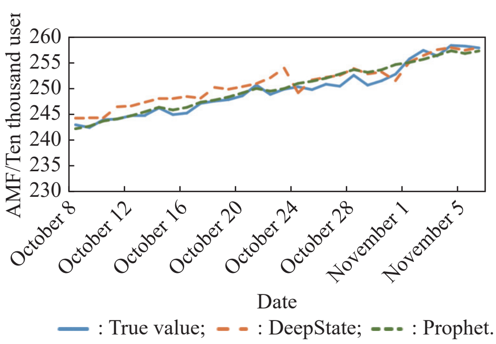

Fig. 6 Prediction of AMF capacity

Table 2 Accuracy evaluation of AMF capacity

Parameter	Prophet	DeepState
MSE	0.000029	0.008465
R^2	0.944	0.83

4.2.2 SMF

The true value, the predicted value using Prophet algorithm, and the predicted value using DeepState algorithm for the registered users of SMF are shown in Fig. 7. The results of R^2 and MSE to evaluate the accuracy are depicted in Table 3.

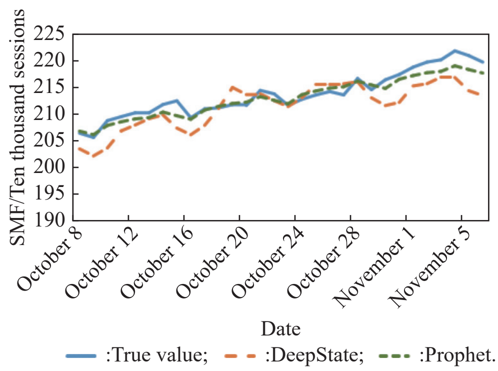

Fig. 7 Prediction of SMF capacity

Table 3 Accuracy evaluation of SMF capacity

Parameter	Prophet	DeepState
MSE	0.000068	0.048908
R^2	0.893	0.36

4.2.3 UDM

The true value, the predicted value using Prophet algorithm, and the predicted value using DeepState algorithm for the authorization users of UDM are shown in Fig. 8. The results of R^2 and MSE to evaluate the accuracy are depicted in Table 4.

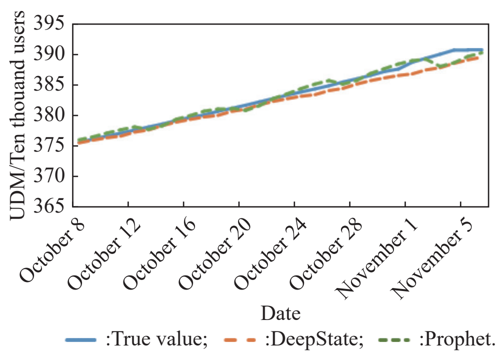

Fig. 8 Prediction of UDM capacity

Table 4 Accuracy evaluation of UDM capacity

Parameter	Prophet	DeepState
MSE	0.000014	0.002061
R^2	0.954	0.98

From the graphic results of three network elements, it can be seen that the Prophet algorithm performs better for predicting AMF and SMF, while the performance is similar for predicting UDM.

For the results of MSE, using Prophet, the performance is always better than using DeepState. For the results of R^2, the prediction of AMF and SMF is better using Prophet, while the R^2 is quite similar for predicting UDM. Moreover, the authorization users of UDM are normally quite stable. Although AMF is related to the mobility of users, the number of registered users of AMF in one province is also stable. The PDU sessions of SMF are related to the users using the business, which are in a dynamic state, therefore, compared to AMF and UDM, there is fluctuation between the predicted values for SMF.

5. Conclusions

The architecture of STIN is depicted in this paper. Three stages of integration and evolution routes are described in detail. As an important network in STIN, the response efficiency of core networks expansion construction should be improved. The method of time series predication is innovatively applied to the core network in this paper. Based on the correlation analysis between each feature and each network elements capacity, the registered users of AMF, authorization users of UDM, PDU sessions of SMF are predicted in combination with the number of NU, and NT. Taking the data from real network, a good prediction result is observed, with R^2 being 0.954 maximum and MSE being 0.000014 minimum.

References

[1] TANG F L. Dynamically adaptive cooperation transmission among satellite-ground integrated networks. Proc. of the Conference on Computer Communications, 2020: 1559−1568.

[2] DING H X, ZHU S C, MENG S C L, et al. Matching-based resource allocation for satellite-ground network. Sensors, 2022, 22(21): 8436–8447.

[3] YANG L, MA W D, WEI D B. Offloading strategy in satellite-ground network. Journal of Physics: Conference Series, 2022, 2290(1): 012118–012125.

[4] LIU Y, DING S H, HUANG J X, et al. Vertical handover of satellite-ground fusion network based on time and location under early access strategy. Proc. of the International Conference on Smart Computing and Communication, 2022: 384−393.

[5] ZHAO M, CHEN C, LIU L, et al. Orbital collaborative learning in 6G space-air-ground integrated networks. Neurocomputing, 2022, 497: 94–109.

[6] LIU Y, PENG M G. Satellite-ground fusion intelligent networking: vision and key technologies. Journal of Beijing University of Posts and Telecommunications, 2021, 44(6): 1–12. (in Chinese)

[7] GUTIERREZ-ESTEVEZ D M, GRAMAGLIA M, DOMENICO A D, et al. Artificial intelligence for elastic management and orchestration of 5G networks. IEEE Wireless Communications, 2019, 26(5): 134–141.

[8] HAN S, MA H B, ZHANG P, et al. Improved MPEG-4 high-efficiency AAC with variable-length soft-decision decoding of the quantized spectral coefficients. China Communications, 2019, 16(10): 65–82.

[9] CHEN M Z, WEI X, GAO Y, et al. Deep-broad learning system for traffic flow prediction toward 5G cellular wireless network. Proc. of International Wireless Communications and Mobile Computing, 2020: 940−945.

[10] HAN S, WANG Z L, WANG G Q, et al. Automatic association of cross-domain network topology. Proc. of the IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 2022: 1173−1178.

[11] NIE L S, NING Z L, OBAIDAT M S, et al. A reinforcement learning-based network traffic prediction mechanism in intelligent internet of things. IEEE Trans. on Industrial Informatics, 2021, 17(3): 2169–2180.

[12] SELVAMANJU E, SHALINI V B. Machine learning based mobile data traffic prediction in 5G cellular networks. Proc. of the 5th International Conference on Electronics, Communication and Aerospace Technology, 2021: 1318−1324.

[13] HAN S, MA H B, CHEN D, et al. Streaming video optimization in mobile communications. Proc. of the IEEE/CIC International Conference on Communications in China, 2018: 495−499.

[14] BARADIE S, REDDY R, LIPPS C, et al. Managing the fifth generation (5G) wireless mobile communication: a machine learning approach for network traffic prediction. Proc. of the IEEE 26th International Symposium on Mobile Communication Technologies and Applications, 2022. https://ieeexplore.ieee.org/abstract/document/9861870.

[15] ALAWE I, KSENTINI A, HADJADJ-AOUL Y, et al. Improving traffic forecasting for 5G core network scalability: a machine learning approach. IEEE Network, 2018, 32(6): 42–49.

[16] DENG D S. Research on Anomaly Detection method based on DBSCAN clustering algorithm. Proc. of the 5th International Conference on Information Science, Computer Technology and Transportation, 2020: 439−442.

[17] HIRAI S, BABA H, MATSUMOTO M, et al. Machine learning based performance prediction for cloud-native 5G mobile core network. Proc. of the IEEE Wireless Communications and Networking Conference, 2022: 1230−1235.

[18] CHAKRABORTY P, CORICI M, MAGEDANZ T. A comparative study for time series forecasting within software 5G networks. Prof. of the 14th International Conference on Signal Processing and Communication Systems, 2020. DOI: 10.1109/ICSPCS50536.2020.9310033.

[19] CHAKRABORTY P, CORICI M, MAGEDANZ T. System failure prediction within software 5G core networks using time series forecasting. Proc. of the IEEE International Conference on Communications Workshops, 2021. DOI: 10.1109/ICCWorkshops50388.2021.9473530.

[20] KURANAGE M P J, NUAYMI L, BOUABDALLAH A, et al. Deep learning based resource forecasting for 5G core net-

work scaling in kubernetes environment. Proc. of the 8th International Conference on Network Softwarization, 2022: 139−144.

[21] NGUYEN H T, DO T V, ROTTER C. Scaling UPF instances in 5G/6G core with deep reinforcement learning. IEEE Access, 2021, 9: 165892–165906.

[22] YAO H P, WANG L Y, WANG X D, et al. The space-terrestrial integrated network: an overview. IEEE Communications Magazine, 2018, 56(9): 178–185.

[23] CHEN L, TANG F L, LI Z T, et al. Time-varying resource graph based resource model for space-terrestrial integrated networks. Proc. of the Conference on Computer Communications, 2021. DOI: 10.1109/INFOCOM42981.2021.9488855.

[24] ZHANG N, ZHANG S, YANG P, et al. Software defined space-air-ground integrated vehicular networks: challenges and solutions. IEEE Communications Magazine, 2017, 55(7): 101–109.

[25] LIU J J, SHI Y P, FADLULLAH Z M, et al. Space-air-ground integrated network: a survey. IEEE Communications Surveys and Tutorials, 2018, 20(4): 2714–2741.

[26] ARANI A H, HU P, ZHU Y Y. Fairness-aware link optimization for space-terrestrial integrated networks: a reinforcement learning framework. IEEE Access, 2021, 9: 77624–77636.

[27] WEI L H, SHUAI J C, LIU Y, et al. Service customized space-air-ground integrated network for immersive media: architecture, key technologies, and prospects. China Communications, 2022, 19(1): 1–13.

[28] WANG X Y, SUN T, DUAN X D, et al. Holistic service-based architecture for space-air-ground integrated network for 5G-advanced and beyond. China Communications, 2022, 19(1): 14–28.

[29] ZHOU Z H, MA J X, WAN G N, et al. Research on 6G satellite-terrestrial integrated network architecture. Proc. of the 7th International Conference on Cyber Security and Information Engineering, 2022: 1057−1062.

[30] DONG G, PAN L, ZHANG Y, et al. Air-space-ground integrated information network: technology, development and prospect. Proc. of the 8th International Conference on Signal and Information Processing, Networking and Computers, 2022: 883−891.

[31] SHENG J, CAI X Q, LI Q Y, et al. Space-air-ground integrated network development and applications in high-speed railways: a survey. IEEE Trans. on Intelligent Transportation Systems, 2021, 23(8): 10066–10085.

[32] PEI Y S, MIAO S Y, LI F C. Challenges and opportunities for terrestrial network operators in the air-space-ground integrated network. Proc. of the International Conference on Parallel and Distributed Processing with Applications, Big Data and Cloud Computing, Sustainable Computing and Communications, Social Computing and Networking, 2020: 1400−1404.

[33] CHI L H, LIN C W, LIN W H, et al. Research on development of space-ground integration information network. Proc. of the International Conference on Urban Engineering and Management Science, 2020: 29−32.

[34] PEI Y, ZHANG Z, WANG T. Development vision and challenge of air-space-ground integrated network. Designing Techniques of Posts and Telecommunications, 2020, 63(4): 15–20.

[35] TAYLOR S J, LETHAM B. Forecasting at scale. The American Statistician, 2018, 72(1): 37–45.

[36] CHAKRABORTY P, CORICI M, MAGEDANZ T. A comparative study for time series forecasting within software 5G networks. Proc. of the 14th International Conference on Signal Processing and Communication Systems, 2020. DOI: 10.1109/ICSPCS50536.2020.9310033.

[37] RANGAPURAM S S, SEEGER M W, GASTHAUS J, et al. Deep state space models for time series forecasting. Proc. of the 32nd Conference on Neural Information Processing Systems, 2018: 7796−7805.

[38] LIU Z X, FENG Y, YIN Y N, et al. DeepState: selecting test suites to enhance the robustness of recurrent neural networks. Proc. of the 44th International Conference on Software Engineering, 2022: 598−609.

[39] ZHANG D B. A coefficient of determination for generalized linear models. The American Statistician, 2017, 71(4): 310–316.

[40] ALMOHAMMEDI A M, ZERGUINE A, DERICHE M, et al. Artificial bee colony DLMS beyond mean square error boundary in ad-hoc WSN. Proc. of the 14th International Conference on Computational Intelligence and Communication Networks, 2022: 572−576.

[41] YAMAKI S, SEKI S, SUGITA N, et al. Performance evaluation of cross correlation functions based on correlation filters. Proc. of the 20th International Symposium on Communications and Information Technologies, 2021: 145−149.

Biographies

HAN Sai was born in 1988. She received her Ph.D. degree from Technical University of Brausch-weig in 2016. She is a senior engineer in China Unicom Research Institute. Her research interests include autonomous network and network intelligence.
E-mail: hans29@chinaunicom.cn

LI Ao was born in 1994. She received her M.S. degree from University of Manchester in 2017. She is an engineer in China Unicom Research Institute. Her research interests include autonomous network and network intelligence.
E-mail: lia12@chinaunicom.cn

ZHANG Dongyue was born in 1995. She received her M.S. degree from Beijing Forestry University in 2019. She is an engineer in China Unicom Research Institute. Her research interests include autonomous network and network intelligence.
E-mail: zhangdy131@chinaunicom.cn

ZHU Bin was born in 1980. He received his B.S. degree from Beijing University of Posts and Telecommunications in 2006. He is an engineer in China Unicom Research Institute. His research interests include key issues of satellite-terrestrial integrated networks.
E-mail: zhubin15@chinaunicom.cn

WANG Zelin was born in 1983. He received his M.S. degree from Northwestern Polytechnical University in 2008. He is a senior engineer in China Unicom Research Institute. His research interests include Internet protocol (IP), cloud-network, whitebox, and software defined network.
E-mail: wangzl172@chinaunicom.cn

WANG Guangquan was born in 1968. He received his B.S. degree from Nanjing University of posts and Telecommunications in 1989. He is a professorate senior engineer in China Unicom Research Institute. His research interests include network planning, new technology evolution, and standard establishment.
E-mail: wanggq122@chinaunicom.cn

MIAO Jie was born in 1983. He received his Ph.D. degree from Beijing University of Posts and Telecommunications. He is a senior engineer in China United Network Communications Group Company Limited. His research interests include network function virtualization, 5G network slice, and automatic support system.
E-mail: miaojie9@chinaunicom.cn

MA Hongbing was born in 1967. He received his M.S. degree from Beijing University of posts and Telecommunications in 1989. He is a professorate senior engineer in China United Network Communications Group Company Limited. His research interests include new technology research in wireless communication, standard establishment, and technical testing.
E-mail: mahb@chinaunicom.cn