Original Research Article

AI-powered solutions for network traffic optimization

Ming Tang

Communication University of China, Nanjing, Nanjing, Jiangsu, 211172, China

Abstract: With the explosion and increasing complexity of network traffic, management faces various technical challenges. Artificial intelligence (AI) technologies, including machine learning, deep learning, reinforcement learning and big data analysis, are used for traffic feature analysis, prediction, dynamic scheduling and optimization strategy formulation. The proposed optimization scheme is based on these techniques, including architecture construction and data processing.

Keywords: Network traffic; Artificial intelligence; Traffic optimization; Machine learning; Deep learning

1. Introduction

With the rapid growth of internet traffic and complex applications like video streaming and cloud gaming, traditional management methods face challenges. AI offers new opportunities for optimizing network traffic with advanced data processing and analysis.

2. Analysis of current network traffic and challenges

(1) Growth trend of network traffic

The rise of internet adoption and smart devices has caused a surge in traffic. Omdia predicts mobile data traffic will double in the next five years with 5G expansion. By 2028, 5G and FWA penetration will reach 52%. Apps like video streaming and cloud gaming will further strain network infrastructure.

(2) Complexity of network traffic composition

Network traffic now includes a mix of traditional services like browsing and email, alongside newer services such as video streaming, cloud gaming, and social media. Platforms like Netflix and Tencent Video generate high traffic, while IoT devices further complicate traffic management with constant data transmission.

(3) Technical challenges in network traffic management

Network traffic management faces challenges like transmission delays, data loss, and noisy data requiring filtering[1]. It demands significant storage and computing resources, often leading to shortages. As sensitive data grows, robust security and scalability are crucial to avoid performance degradation as traffic increases.

3. Application principles of artificial intelligence technology in network traffic optimization

(1) Machine learning algorithms in traffic feature analysis

Machine learning algorithms extract major features from large-scale network traffic data, supervise learning algorithms (such as decision tree), and classify traffic using labeled data (such as source IP, elephant IP, port, protocol). In the dataset containing 50,000 records, the decision trees reached about 90% of the classification accuracy. Unsupervised clustering algorithms such as K-Means group traffic data through similarity, identify traffic fluctuations and other patterns, and improve traffic management[2].

(2) Deep learning models in traffic prediction

Deep learning models are good at time-series data processing, and are a general feature of network traffic. Long-term and short-term memory (LSTM) networks capture long-term dependencies and short-term fluctuations. For example, the LSTM model using traffic data from the past month and a 30-minute time step. Predict the next 12 hours of traffic to help allocate network resources.

(3) Reinforcement learning in dynamic traffic scheduling

Reinforcement learning through agent-environment interaction optimization strategy. In traffic scheduling, the agent (traffic scheduler) can accept real-time conditions, such as link bandwidth and traffic requirements [3]. When traffic surges, reinforcement learning dynamically adjusts the routing and bandwidth. Repeat 200 times in a network containing 20 nodes and 50 links to reduce congestion and improve efficiency.

(4) Big data analytics in traffic optimization strategy

Extensive data analysis identifies traffic patterns by merging multi-source traffic data (including user behavior, device status, and services). For example, by analyzing the video viewing data of 1,000 users, you can find that the peak HD video traffic is three times the standard definition, providing a strategy to optimize bandwidth and improve resource utilization.

4. Implementation of AI-based network traffic optimization solution

(1) Architecture construction of the solution

At the data collection layer, using protocols such as the simple network management protocol (SNMP) and NetFlow), traffic data is collected from network devices (such as routers and switches), including source IP, elephant IP, port number, traffic size, and transmission time. Data were collected every 5 min. Ensure the real-time performance and integrity of the data. At the data processing layer, preprocessing collected data (such as Hadoop, vs. Spark), clearing duplicate data, filling in missing values, converting data formats, etc.) to improve data quality and meet subsequent analysis requirements[4]. The intelligent decision-making layer uses machine learning and deep learning algorithms such as convolutional neural network (CNN) and long-term short-term memory network (LSTM) to analyze and process data, predict traffic trends, and formulate traffic optimization strategies[5]. The application program interface layer provides the visual interface and the API interface. Administrators can view network traffic status, adjust optimization parameters, and manage traffic tasks.

(2) Data collection and processing

During the data collection phase, in addition to traffic data from network devices, user behavior data—such as website visits, browsing time, and downloaded content—must also be gathered[6]. Device status information, including CPU and memory utilization, as well as application service data, is also necessary. In medium-sized enterprise networks, daily traffic data can reach several gigabytes[7]. Distributed storage technology ensures sufficient storage capacity and quick data access. During data processing, mining algorithms, like a priori algorithms, identify correlations between data types, such as the link between video app usage and peak traffic, which aids in optimizing future traffic management strategies.

(3) Algorithm model deployment

Machine learning algorithms, like Support Vector Machines (SVM), are used to classify normal and abnormal traffic, identifying potential network attacks or anomalies. The model, trained on 10,000 normal and 2,000 abnormal records, achieves around 95% accuracy. Deep learning models, such as Recurrent Neural

Networks (RNN), predict traffic flow based on past data, with prediction errors controlled within a specific range[8]. Additionally, advanced algorithms like Q-learning enable dynamic adjustments of traffic routes and bandwidth allocation based on real-time network conditions, optimizing transmission efficiency.

(4)System integration and testing

In the testing phase, network simulation tools are employed to create a virtual environment that mimics the actual network topology, incorporating various traffic patterns such as burst peaks and sustained high traffic[9]. The optimized solution's performance is evaluated under different conditions. A stress test examines the system's stability and reliability by progressively increasing the traffic load to assess the maximum capacity[10]. For instance, when traffic reaches five times the normal level, the system maintains response speed and data accuracy. The architecture of the AI-based solution is shown in the **Figure 1** below.

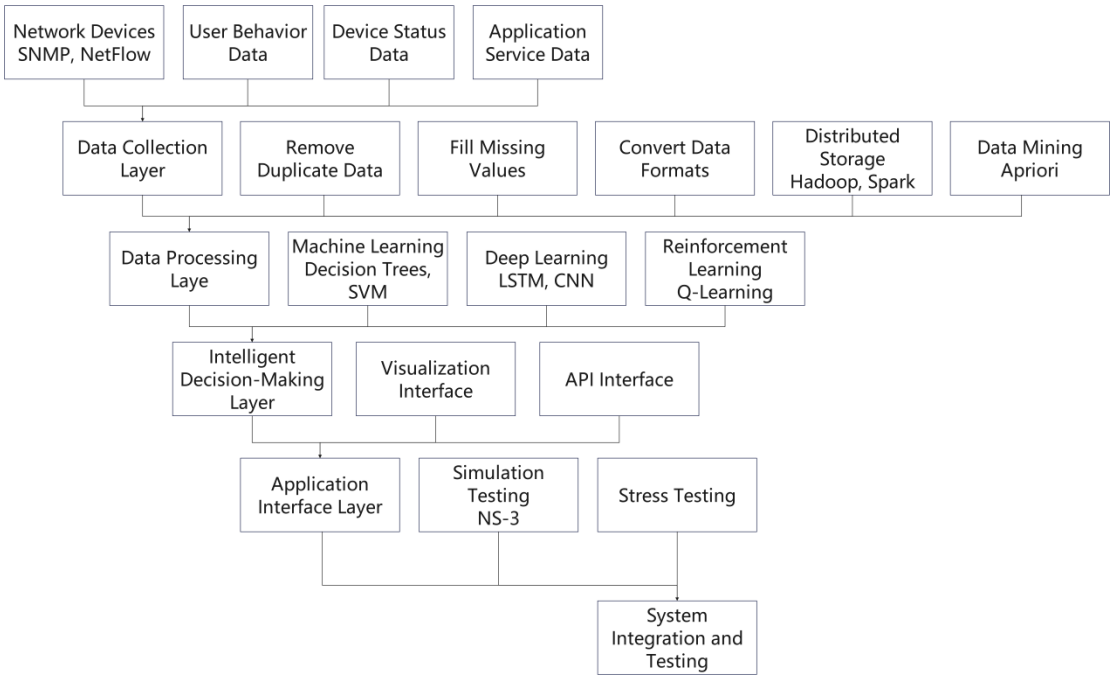

Figure 1. Architecture of AI-based network traffic optimization solution.

5. Conclusion

AI-driven network traffic optimization offers significant improvements in managing complex and growing network demands. By leveraging machine learning, deep learning, reinforcement learning, and big data analytics, future networks can achieve greater efficiency, scalability, and reliability, paving the way for enhanced performance in emerging applications and technologies.

Author the author

Ming Tang, born in February 1973, male, native of Dongyang, Zhejiang Province, Han nationality, bachelor's degree, senior engineer, research direction: computer network, high-performance computing.

References

[1] Teng Y .Research on the intelligent detection of network marketing error flow based on artificial intelligence. International Journal of High-speed Electronics and Systems, 2025, (Pre-published):

[2] Xu S, Wang Q. Application of artificial intelligence in network traffic analysis and prediction. Progress in

Computer, Signal, and Systems, 2024,8 (5):

[3] Aceto G, Gipolo F, Guida C, et al. Using a generative artificial intelligence model to generate synthetic and privacy-protected traffic tracking and train a network intrusion detection system [J]. Journal of Networks and Computer Applications, 2024,229103926-103926.

[4] Retraction article: Retraction: traffic forecasting and transmission scheduling of AI-driven cognitive wireless networks [J]. The International Journal of Computers and Applications, 2022, page 44 (6): page 596-596.

[5] Jiyuan R, Yunhou Z, Zhe W, et al. AI-based network traffic analysis and automatic optimization technology. Mathematics, Biological Sciences and Engineering: Bachelor of Science Medal, 2022,19 (2): 1775-1785.

[6] Artificial intelligence; research results of Nanjing University of Posts and Telecommunications (multidimensional traffic control inspired by artificial intelligence) [J]. The Journal of Robotics and Machine Learning, 2019,

[7] Mathematics; research results report from the Institute of Artificial Intelligence in the field of mathematics (deep reinforcement learning for router selection in networks with large traffic) [J]. WebWeekly News, 2019,

[8] Artificial Intelligence; researchers from the University of Heroliat report findings of artificial intelligence (AI manages network traffic on 5g wireless networks) [J]. The Journal of Robotics and Machine Learning, 2019,

[9] Arafah M ,Phillips I ,Adnane A , et al.An enhanced BiGAN architecture for network intrusion detection[J].Knowledge-Based Systems,2025,314113178-113178.

[10] Bistene V J ,Santos D P F A ,Chagas D E C , et al.Modeling Network Traffic Generators for Cyber Ranges: A Systematic Literature Review[J].Journal of Network and Systems Management,2025,33(2):35-35.